Microbiota-Host Symbiosis in First-Onset Pediatric Inflammatory Bowel Disease

Walid Abd El-Fattah El-Sayed Mottawea

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements
for the Doctorate of Philosophy degree in Microbiology and Immunology

Department of Biochemistry, Microbiology and Immunology
Faculty of Medicine
University of Ottawa

Abstract

In recent years, the association between inflammatory bowel diseases (IBDs) and gut microbiota has been extensively studied in adults using post-treatment cohorts of patients. However, microbial composition and functional interplay between host genetics and microorganisms in newly diagnosed early IBD onset remain poorly defined. Using colonoscopic mucosal washes to collect mucosal-luminal microbiota from different intestinal locations, we studied the gut microbiome in a large number of children with either Crohn's disease (CD) or ulcerative colitis (UC). Although no significant difference in the diversity was evident between the gut microbiota of IBD-affected and control children, the microbiome of IBD subjects is characterized by an increased abundance of potent hydrogen sulfide (H_2S) producers and decreased abundance of beneficial butyrate producers. Microbiota and proteomic profiling revealed that the abundance of *Atopobium parvulum*, a potent H_2S producer, was associated with increased CD severity and a concurrent reduction in the expression of the host H_2S detoxification pathway. Gnotobiotic and conventionalized colitis-susceptible interleukin-10-deficient (*Il-10*$^{-/-}$) mice showed that *A. parvulum* induces severe colitis, a phenotype requiring the presence of the gut microbiota. In addition, administration of bismuth, an H_2S scavenger, prevented *A. parvulum*-induced colitis in *Il-10*$^{-/-}$ mice. Our findings have identified *A. parvulum* as a major mediator of inflammation severity. We also reveal an imbalance between the H_2S production and detoxification in the gastrointestinal tract of pediatric IBD patients. Altogether, our findings provide new avenues for diagnostics as well as therapies to treat IBD.

Acknowledgements

I would like to thank all persons who helped me and supported me during my work in this research project and writing this thesis.

First, I would like to thank my supervisor, Dr. Alain Stintzi, for his constant invaluable support, advices and thoughts over the past six years. I owe him a huge debt of gratitude for his precious expert guidance and encouragement since I started work in his lab as a PhD student with little knowledge in the field of microbiome and high throughput analysis.

Next, a massive thank you to my wife, Salma, and my lovely children, Jana, Nour, and Zeyad for putting up with my moods and stress, and for being away from family times during my lab work and writing this thesis, and to you this thesis is dedicated. I am also forever grateful to my mum and dad for all they sacrificed for me and for being away from them over the past 6 years.

My great thanks to the inflammatory bowel disease research team at Children Hospital of Eastern Ontario (CHEO), particularly Dr. David Mack and Ruth Singleton for their great efforts in collecting the biological samples involved in this work and for spending their time to keep me updated with samples-related information.

I would also like to thank the members of my thesis advisory committee; Dr. David Mack and Dr. Marc Andre Langlois; for their invaluable thoughts, feedbacks, discussions, comments and directions during my work in this research project.

A special thank you to my friend, Turki Abujamel, for his assistance, friendly cooperation and valuable discussions in different aspects of this research. I would also like to thank Dr. Martin Stahl and Jennifer Li for proofreading my thesis writing.

To my lab-mates, Dr. Ibrahim Taher, Dr. Olle de Bruin, Dr. Martin Stahl, Dr. Christina Xueqi Wang, Turki Abujamel, Momen Askora, Annika Flint, James Butcher, M.D. Kelly Grzywacz, Zack Li, Guillaume Romain, Jennifer Li and Hai Nguyen, thank you for your wonderful friendship, continual support and warm humour and for everything you do to make life easier for me.

Table of Contents

Abstract ii

Acknowledgements iii

Table of Contents iv

List of Abbreviations xii

Glossary of Terms xvi

List of Figures xvii

List of Tables xxi

CHAPTER 1. Introduction 1

1.1. Clinical manifestations of pediatric IBD. 1

1.2. Epidemiology of IBD. 2

1.3. Etiopathogenesis of IBD. 3

1.3.1. Immunology of IBD. 3

1.3.2. Genetic susceptibility association with IBD. 5

1.3.3. Environmental risk factors in IBD. 6

1.4. Normal gut microbiota. 10

1.5. Roles of gut microbiota in human health. 12

1.6. Gut microbiota dysbiosis in IBD. 15

1.7. Different approaches for gut microbiota characterization. 19

1.7.1. 16S rRNA-based microarray. 20

1.7.2. 16S rRNA sequencing approaches. 21

1.8. Dysfunctionality of gut microbiota in IBD. 24

1.8.1. Hydrogen sulfide metabolism in IBD. 25

1.8.2. Butyrate metabolism in IBD. 30

1.9. Hypothesis and objectives of this study. 34

CHAPTER 2. Materials and Methods 35

2.1. Ethics statement. 35

2.2. Subject selection and sampling. 35

2.3. Metagenomic DNA extraction. 37

2.4. 16S rRNA-V6 454-pyrosequencing. 37

2.5. 454-pyrosequencing data analysis. 38

2.6. 16S rRNA-V6 library construction for Illumina sequencing. 39

2.7. Illumina sequencing data analysis. 40

2.8. Statistical analysis. 41

2.9. Prediction of the gut microbial functionalities in control, CD and UC groups. 41

2.10. Quantitative polymerase chain reaction (qPCR)-validation of sequencing results. ...42

2.11. Dextran Sodium Sulfate (DSS)-induced acute colitis model. 43

2.12. Assessment of cytokines in the intestinal tissues of DSS-treated mice. 44

2.13. *Il-10*$^{-/-}$ Mice experiments and tissue processing. 45

2.14. Mouse endoscopy. 46

2.15. Real time RT-PCR on mouse intestinal samples. 46

2.16. Statistical analyses of *Il-10*$^{-/-}$ mice results. 47

2.17. Proteomic analysis of biopsies using super-SILAC-based quantitative mass spectrometry. 47

2.18. Total RNA extraction from mucosal aspirates for real-time qRT-PCR. 48

2.19. qRT-PCR quantification of hydrogen sulfide detoxification genes expression level. 49

CHAPTER 3.Results 51

3.1. Demographic and clinical characteristics of the pediatric cohort employed. 51

3.2. Differences in microbial signature identified from mucosal aspirates, mucosal biopsies and stools. 53

3.2.1. Diversity of the gut microbiota from different sampling sources. 53

3.2.2. Taxonomic composition of the gut microbiota from biopsies, mucosal aspirates and stool samples. 57

3.3. Microbial diversity and structure along the gut of healthy children. 62

3.3.1. Gut microbial diversity of healthy children. 62

3.3.2. Microbial structure along the length of the intestinal tract of non-IBD children. . 64

3.3.3. The diversity and composition of core microbiota along the intestine of healthy children. 69

3.4. Characterization of left colon (LC) microbiota in pediatric IBD. 71

3.4.1. Microbial diversity of the left colon in pediatric IBD. 71

3.4.2. Dysbiosis of the left colon (LC) microbiota in pediatric IBD. 73

3.4.3. Depletion of the LC core microbiota in pediatric patients with IBD. 76

3.4.4. The left colon microbiota is associated with IBD activity. 79

3.5. Characterization of the right colon (RC) microbiota in pediatric IBD. 81

3.5.1. Microbial diversity of the right colon in pediatric IBD. 81

3.5.2. Dysbiosis of the RC microbiota in pediatric IBD. 83

3.5.3. Depletion of the RC core microbiota in pediatric IBD. 90

3.5.4. The right colon microbial taxa are associated with pediatric IBD activity. 93

3.5.5. qPCR validation of the Illumina sequencing results from RC samples. 95

3.6. Characterization of the gut microbiota at the terminal ileum (TI) of pediatric IBD. ... 97

3.6.1. Diversity of the terminal ileal (TI) microbiota in pediatric IBD. 97

3.6.2. Dysbiosis of the TI microbiota in pediatric IBD. 97

3.6.3. Depletion of the TI core microbiota in pediatric IBD. 102

3.6.4. The terminal ileal microbiota is associated with pediatric IBD activity. 105

3.7. Alteration of the identified CD-specific microbial biomarkers across age, gender and intestinal site. 107

3.8. Functional dysbiosis of gut microbiota along the intestine of new onset pediatric IBD. 109

3.9. *Atopobium parvulum* modulates the severity of pediatric IBD inflammation in presence of other gut microbes. 113

3.9.1. *Atopobium parvulum* induces rapid and severe pan-colitis in 129/SvEv *Il-10*$^{-/-}$ mice under specific pathogen free (SPF) conditions. 113

3.9.2. Bismuth alleviates *A. parvulum*-induced colitis. 116

3.9.3. Gut commensal microbiota are essential for *Atopobium parvulum*-induced colitis development. 118

3.10. Alteration of human colonic metaproteome in pediatric IBD. 121

CHAPTER 4. Discussion 125

4.2. Conclusions: 147

4.3. Future Directions. 151

CHAPTER 5. References.....154

Contribution of Collaborators 179

Appendix I: Primers used for constructing the 16S rRNA-V6 Illumina library. 180

Appendix II: Characteristics of IBD patients and control subjects employed in the current study. 181

Appendix III: Taxa that vary significantly in abundance in at least one of the three intestinal locations of non-IBD adolescents (LC; left colon, RC; right colon and TI; terminal ileum). 196

Appendix IV: Left colon taxa that vary significantly in abundance in at least one of the three age groups of non-IBD adolescents (A;<10 years, B; 11-14 years and C; 15-18 years). 201

Appendix V: Right colon taxa that vary significantly in abundance in at least one of the three age groups of non-IBD adolescents (A;<10 year, B; 11-14 year and C; 15-18 year). 203

Appendix VI: Terminal Ileum taxa that vary significantly in abundance in at least one of the three age groups of non-IBD adolescents (A;<10 years, B; 11-14 years and C; 15-18 years). 206

Appendix VII: Core OTUs (0.75) that vary significantly in abundance in at least one of the three intestinal regions tested in non-IBD adolescents (LC; left colon, RC; right colon and TI; terminal ileum). 207

Appendix VIII: Left colon taxa that vary significantly in abundance in at least one of the three pairwise comparisons performed (controls vs. CD; controls vs. UC; and CD vs. UC). 215

Appendix IX: Left colon core OTUs (0.75) that vary significantly in abundance in at least one of the three pairwise comparisons performed (controls vs. CD; controls vs. UC; and CD vs. UC). 218

Appendix X: Left colon taxa that vary significantly in abundance in CD patients in at least one of the three pairwise comparisons performed (mild vs. moderate; mild vs. severe; and moderate vs. severe). 222

Appendix XI: Left colon taxa that vary significantly in abundance in UC patients in at least one of the three pairwise comparisons performed (mild vs. moderate; mild vs. severe; and moderate vs. severe). 224

Appendix XII: Right colon taxa that vary significantly in abundance in at least one of the three pairwise comparisons performed (controls vs. CD; controls vs. UC; and CD vs. UC). 227

Appendix XIII: Right colon core OTUs (0.75) that vary significantly in abundance in at least one of the three pairwise comparisons performed (controls vs. CD; controls vs. UC; and CD vs. UC).....231

Appendix XIV: Right colon taxa that vary significantly in abundance in CD patients in at least one of the three pairwise comparisons performed (mild vs. moderate; mild vs. severe; and moderate vs. severe).....233

Appendix XV: Right colon taxa that vary significantly in abundance in UC patients in at least one of the three pairwise comparisons performed (mild vs. moderate; mild vs. severe; and moderate vs. severe).....236

Appendix XVI: Terminal Ileum taxa that vary significantly in abundance in at least one of the three pairwise comparisons performed (controls vs. CD; controls vs. UC; and CD vs. UC).....238

Appendix XVII: Terminal ileum core OTUs (0.75) that vary significantly in abundance in at least one of the three pairwise comparisons performed (controls vs. CD; controls vs. UC; and CD vs. UC).....242

Appendix XVIII: Terminal ileum taxa that vary significantly in abundance in CD patients in at least one of the three pairwise comparisons performed (mild vs. moderate; mild vs. severe; and moderate vs. severe).....244

Appendix XIX: Terminal ileum taxa that vary significantly in abundance in UC patients in at least one of the three pairwise comparisons performed (mild vs. moderate; mild vs. severe; and moderate vs. severe).....247

Appendix XX: Microbial metabolic pathways with significant differential abundance in the LC communities of new onset pediatric IBD patients. 248

Appendix XXI: Microbial metabolic pathways with significant differential abundance in the RC communities of new onset pediatric IBD patients. 250

Appendix XXII: Microbial metabolic pathways with significant differential abundance in the TI communities of new onset pediatric IBD patients.....255

Appendix XXIII: Taxa that vary significantly in abundance in Il-10-/- mice in response to *A. parvulum* colonization and/or bismuth administration. 257

Appendix XXIV: List of all differentially expressed proteins and their variable importance in projection scores (VIP) derived from the calculated PLS-DA model. 268

Appendix XXV: List of differentially expressed mitochondrial proteins and their variable importance in projection scores (VIP) derived from the calculated PLS-DA model.....297

Curriculum Vitae----------303

List of Abbreviations

AMP	Adenosine Monophosphate.
ATG16L1	Autophagy Like Protein 16-1.
BCON	Colon Biopsy from Normal Tissue.
BLAST	Basic Local Alignment Search Tool.
BMI	Body Mass Index.
BP-FAT	Biological Process Functional Annotation.
CARD15	Caspase Recruitment Domain-Containing Protein 15.
CCR6J	CC Chemokine Receptor 6.
CD	Crohn's Disease.
CFU	Colony Forming Unit.
COasp	Colon Aspirate.
COX4-1	Cytochrome c Oxidase Subunite 4 Isoform 1.
Ct	Threshold Cycle.
DC	Dendridic Cells.
DEPC	Diethylpyrocarbonate.
DLG5	Disks Large Homolog 5.
DMBT1	Deleted in Malignant Brain Tumors 1.
DSR	Dissimilatory Sulfite Reductase.
DSS	Dextran Sodium Sulfate.
ETHE-1	Ethylmalonic Encephalopathy 1.
FAB	Fastidious Anaerobic Broth.

FISH	Flourescent *In-Situ* Hybridization.
FLASh	Fast Length Adjustment of Short Reads.
GALT	Gut Associated Lymphoid Tisuue.
GF	Germ Free.
GPR	G-Protein Coupled Receptor.
GWAS	Genome-Wide Association Study.
H_2S	Hydrogen Sulfide.
HDAC	Histone Deacetylase.
hGAPDH	Human Glyceraldehyde 3-Phosphate Dehydrogenase.
IBD	Inflammatory Bowel Disease.
IBDU	Inflammatory Bowel Disease type unclassified.
IBS	Irritable Bowel Syndrome.
IL	Interleukin.
INFα	Interferon Alpha.
IRGM	Immunity Related GTPase Family M Protein.
ITLN1	Intelectin 1.
JAK2	Janus Kinase 2.
KEGG	Kyoto Encyclopedia of Genes and Genomes.
LC	Left Colon.
LEfSe	Linear Discriminant Analysis (LDA) Effect Size.
LPS	Lipopolysaccharides.
LTβR-Ig	Lymphotoxin-Beta Receptor-Immunoglobulin Fusion Protein.

MCT-1	Monocarboxylic Acid Transporter 1.
MF-FAT	Molecular Functions Functional Annotation.
NOD2	Nucleotide-Binding Oligomerization Domain-Containing Protein 2.
NSAID	Non-Steroidal Anti-Inflammatory Drugs.
OTU	Operational Taxonomic Unit.
PCA	Principle Component Analysis.
PCDAI	Pediatric Crohn's Disease Activity Index.
PCoA	Principle Coordinate Analysis.
PLS-DA	Partial Least Squares-Descriminant Analysis.
PMN	Polymorphonuclear Leukocyte.
PTGER4	Prostaglandin E receptor 4.
PUCAI	Pediatric Ulcerative Colitis Activity Index.
QIIME	Quantitative Insights Into Microbial Ecology.
RC	Right Colon.
SCFA	Short Chain Fatty Acids.
SILAC	Stable Isotope Labeling by Amino Acids in Cell Culture.
SPF	Specific Pathogen Free.
SQR	Sulfide Quinone Reductase.
SQRDL	Sulfide Quinon Reductase Like Protein.
SRB	Sulfate Reducing Bacteria.
STAT3	Signal Transducer and Activator of Transcription 3.
TGFβ	Transforming Growth Factor Beta.

Th	T Helper Cells.
Thy-1	Thymocyte Differentiation Antigen 1.
TI	Terminal Ileum.
TLR	Toll Like Receptor.
TMT	Thiol Methyl-Transferase.
TNBS	Tri-Nitro-Benzene Sulfonic Acid.
TNFα	Tumor Necrosis Factor Alpha.
T_{Reg}	Regulatory T Cells.
TST	Thiol Sulfur-Transferase.
UC	Ulcerative Colitis.
XBP1	X-box Binding Protein 1.

Glossary of Terms

Alpha-Diversity: (Within-habitat or within-sample diversity) is a number of species (richness) and their proportions (eveness) within a locality or environment (within each sample).

Beta-Diversity: (Between –habitats or among samples diversity) is the changes in species composition (species turnover) among habitats or samples.

Core microbiota: Bacterial taxa that are conserved between individuals.

Dysbiosis: Alteration of the normal gut microbiota composition that affects the symbiotic relationship between the gut microbes and their host.

Enterotype: (Gut type) is a certain configuration of the gut microbiota species that is common between different environments or individuals and defined based on the relative abundance of microbial species in these environments.

High-Throughput: (Many at once) High quantity or amount of work to be done in a given time. For instance, high throughput sequencing means large number of sequences generated at the same experiment or the same run instead of generating a single sequence one after another.

Intestinal homeostasis: is to maintain the stability of the gut environment in response to a change in the external factors.

Microbiota: A population of microorganisms that live in a particular site, habitat or environment including human body or organ.

Mutualism: a symbiotic ecological interactions between two organisms of different species that is beneficial to both organisms.

-omic: a suffix refers to a large-scale study of the totality of something such as genes (genomics), or protein (proteomics).

Relative abundance: The proportion of certain bacterial taxon in the whole microbial community which is calculated by dividing the number of the 16S rRNA reads assigned to that taxon by the total number of reads generated from one sample.

List of Figures

Figure 1.1: Schematic representation of the conserved and hypervariable regions of 16S rRNA gene. **23**

Figure 3.2.1: Diversity of the gut microbiota identified from mucosal aspirates (COasp), mucosal biopsies (BCOA for inflamed tissues and BCON for non-inflamed tissues) and stool samples. **54**

Figure 3.2.2: Diversity of the core gut microbiota (0.75) identified from mucosal aspirates (COasp), mucosal biopsies from inflamed (BCOA) and non-inflamed tissues (BCON) and stool samples. **56**

Figure 3.2.3: Relative abundance of bacterial phyla from colon biopsies, colon aspirates and stool samples. **58**

Figure 3.2.4: Bacterial classes and families that vary significantly in abundance in at least one of the three sample types (Biopsies, stool and aspirates). **60**

Figure 3.2.5: Bacterial genera that vary significantly in abundance in at least one of the three sample types (Biopsies stool and aspirates). **61**

Figure 3.3.1: Gut microbiota exhibit similar diversity along the intestinal tract of non-IBD children and between different age and gender groups. **63**

Figure 3.3.2: Phylogenetic beta-diversity analyses revealed individuality of microbial communities along the intestinal tract between different subjects. **65**

Figure 3.3.3: Microbial structure along the length of the intestinal tract of healthy children. **66**

Figure 3.3.4: Microbial structure along the length of the intestinal tract of healthy children at different ages. **68**

Figure 3.3.5: The left and right colon of non-IBD children harbour a less diverse core microbiome as compared to the terminal ileum. **70**

Figure 3.4.1: The gut microbiota exhibit a similar diversity at the left colon of pediatric patients with IBD as compared to control children. **72**

Figure 3.4.2: Taxonomic composition of the left colon microbial phyla of IBD patients and control subjects. **74**

Figure 3.4.3: Dysbiosis of the left colon microbiota in pediatric IBD. **75**

Figure 3.4.4: Phylogenetic tree of the microbial taxa detected in the left colon of at least 75% of the samples within each patient cohort. **77**

Figure 3.4.5: CD and UC microbial communities are characterized by a smaller and specific left colon core microbiome as compared to controls. **78**

Figure 3.4.6: Partial least squares analysis of the gut microbiota revealed biomarkers of disease severity. **80**

Figure 3.5.1: The gut microbiota exhibit similar diversity at the right colon of pediatric patients with IBD in comparison to control children based on Illumina HiSeq 2500 and 454-pyrosequencing.. ..... **82**

Figure 3.5.2: Taxonomic composition of the right colon microbiota of IBD patients and control subjects. **84**

Figure 3.5.3: CD-specific microbial biomarkers identified at the right colon.. ..... **85**

Figure 3.5.4: UC-specific microbial biomarkers identified at the right colon. **87**

Figure 3.5.5: Dysbiosis of the right colon microbiota in pediatric IBD. **88**

Figure 3.5.6: Taxonomic composition of the RC microbiota of IBD patients and control subjects based on 16S rRNA pyrosequencing reads. **89**

Figure 3.5.7: Phylogenetic tree of the microbial taxa detected in the right colon of at least 75% of the samples within each patient cohort.. ..... **91**

Figure 3.5.8: The CD and UC microbial communities are characterized by a smaller, specific right colon core microbiota as compared to controls. **92**

Figure 3.5.9: Partial least squares analysis of the right colon microbiota revealed biomarkers of disease severity. **94**

Figure 3.5.10: *Atopobium parvulum* and *Veillonella* are biomarkers of CD and UC severity, respectively, as revealed by qPCR.. ..... **96**

Figure 3.6.1: The gut microbiota exhibit similar diversity at the terminal ileum of pediatric patients with IBD in comparison to control children. **98**

Figure 3.6.2: Taxonomic composition of the terminal ileum microbiota of IBD patients and control subjects. **99**

Figure 3.6.3: Dysbiosis of the terminal ileum microbiota in pediatric IBD. **101**

Figure 3.6.4: Phylogenetic tree of the microbial taxa detected in the terminal ileum of at least 75% of the examined samples within each patient cohort. **103**

Figure 3.6.5: CD and UC are characterized by a depleted terminal ileum core microbiota as compared to controls. **104**

Figure 3.6.6: Partial least squares analysis of the terminal ileum microbiota revealed biomarkers of disease severity. **106**

Figure 3.7.1: Alteration of the CD specific microbiota as a function of age, gender or intestinal location. **108**

Figure 3.8.1: Microbial metabolic pathways with significant differential abundance in the LC (A) and TI (B) communities of new onset pediatric IBD patients. **111**

Figure 3.8.2: Microbial metabolic pathways with significant differential abundance in the RC communities of new onset pediatric IBD patients. **112**

Figure 3.9.1: *Atopobium parvulum* induces rapid and severe pan-colitis in *Il-10*$^{-/-}$ mice. ... **114**

Figure 3.9.2: *A. parvulum* induces cytokine expression in conventionalized *Il-10*$^{-/-}$ mice. . **115**

Figure 3.9.3: Bismuth prevents *Atopobium*-induced colitis. **117**

Figure 3.9.4: *Il-10*$^{-/-}$ mice mono-associated with *A. parvulum* exhibit increased hyperplasia scores and GALT foci and bismuth decreases *A. parvulum* colonization level. **119**

Figure 3.9.5: Gnotobiotic mice mono-associated with *A. parvulum* exhibit an incomplete picture of inflammation post DSS treatment. **120**

Figure 3.10.1: Top representative Gene Ontology (GO) and KEGG pathways enriched in CD patients as compared to control subjects. **122**

Figure 3.10.2: PLS-DA analysis of the host proteome identified the sulfide detoxification pathway as a functional biomarker of Crohn's disease. **123**

Figure 4.1: Schematic representation of the microbiota distribution across the gut lumen and different mucus layers of the large intestine. **128**

Figure 4.2: Model of mitochondrial H2S catabolism. **145**

Figure 4.3: Proposed host-microbe interaction in pediatric IBD. **150**

List of Tables

Table 2.1: Primers used in qPCR validation of sequencing results. **43**

Table 2.2: Primers used in qRT-PCR.....**50**

Table 3.1: Metadata statistics for the subjects involved in the current study.....**52**

CHAPTER 1. Introduction

1.1. Clinical manifestations of pediatric IBD.

Inflammatory bowel disease (IBD) is a group of medical disorders characterized by recurrent inflammation of the gastrointestinal tract. It mainly comprises two different, but related types: ulcerative colitis (UC) and Crohn's disease (CD). However, approximately 10% of IBD cases cannot be definitively distinguished at the time of diagnosis; they are instead categorized as "Inflammatory Bowel Disease type unclassified (IBDU)" [1]. UC inflammation is mainly restricted to the colon with 95% of patients suffering from rectum inflammation with variable degrees of proximal extension. Inflammation mainly affects the mucosal layer and consists of continuous ulcerations, edema and hemorrhaging along the length of the colon [2]. On the other hand, CD is a transmural chronic inflammation of any location within the gastrointestinal tract from the mouth to the perianal region. Inflammed areas are frequently separated by normal mucosa, resulting in characteristic skip areas [2]. Extraintestinal manifestations are also common among 25-40% of IBD cases [3]. These include, but are not limited to, joint arthritis, skin erythema, ocular manifestations, oral ulcerations in CD, fatty liver, metabolic and renal disorders [3]. Children with IBD have a characteristic inflammatory phenotype that is represented by extensive anatomical involvement and rapid progress of the disease within a short period of time. For example, panenteric CD and pancolitic UC are represented in 50.5-61.5% and 70-90% of early-onset IBD cases, respectively, as compared to 42% and 24% for adult patients with CD and UC, respectively [4-7]. Furthermore, the location, symptoms and behaviour of pediatric IBD can vary from the disease in adults. Pediatric IBD mainly affects the large bowel with the terminal ileum and the right colon being the most affected sites in CD with rapid extension to

other anatomical locations over time [4, 5]. Whereas rectal bleeding and diarrhea are the most common symptoms of adult UC and CD, respectively, abdominal pain is the most frequent in childhood cases of IBD [8]. Moreover, 10-40% of childhood IBD cases develop growth retardation, a unique extraintestinal manifestation of pediatric IBD [3, 9]. Finally, the inflammatory behaviour of the disease is predominant in childhood IBD, while in adults, it may be inflammatory, stricturing, penetrating or fistulising [8].

1.2. Epidemiology of IBD.

IBD is distributed worldwide with increasing incidence and prevalence over time [10-12]. The highest reported incidence and prevalence rates are in Canada, Northern Europe and Australia [10, 13-15]. In Canada, the prevalence rate ranges from 161 to 319 per 100,000 persons for CD and 162 to 249 per 100,000 persons for UC, while the incidence rate is 19.2 per 100,000 for UC and 20.2 per 100,000 for CD [13, 14]. In 2012, about 233,000 Canadians have been estimated to have IBD which divides into approximately 129,000 CD and 104,000 UC cases [16]. These estimates indicate that, approximately, 0.67% of the Canadian population suffer from IBD with 10,200 new cases diagnosed annually [16]. Although the highest incidence rates are among the second and the third decades of life [10, 12], 9.61 % of IBD cases diagnosed from 1999 to 2008 are under the age of 18 [12]. Five Canadian provinces (Manitoba, Nova Scotia, Saskatchewan, British Columbia and Alberta) exhibit variable prevalence rates of pediatric IBD that range from 30 to 71 per 100,000 for CD, and 17 to 31 per 100,000 for UC. The pediatric incidence rates vary between 5.4 and 12.0 per 100,000 children for CD, and 3.2 to 5.7 per 100,000 children for UC [14]. Ontario has one of the highest incidence and prevalence rates of early-onset IBD worldwide, and the disease incidence quickly increases in young children [17]. For instance, both the incidence and

prevalence rates of pediatric IBD have markedly increased in Ontario from 9.5 and 42.2 per 100,000 children, respectively, in 1994 to 11.4 and 56.3 per 100,000 children, respectively, in 2005 [17]. While CD is more common than UC in both adults and pediatric IBD [16], childhood IBD affects more males than females, the opposite of what is seen in adult IBD [16, 17]. Furthermore, CD hospitalization is more prevalent at a younger age than with older patients, which is in contrast to what is observed for UC [18]. Finally, genetic associations with IBD have a higher significance in childhood cases compared to adult cases; a positive family history has been confirmed in 25.6% and 26.8% of children compared to 9.77% and 14.2% of adults with CD and UC, respectively [19, 20]. Collectively, these epidemiological measures indicate that early onset IBD is a specific category of IBD.

1.3. Etiopathogenesis of IBD.

The exact etiology of IBD is still unknown, however, it is widely accepted that the pathogenesis results from an alteration of the interaction between gut microbiota and a host immune system with genetic and environmental predispositions [21-23]. Together, the immunological, genetic and environmental architectures of IBD illustrate the role of host microbe interaction in the disease pathogenesis as discussed below.

1.3.1. Immunology of IBD.

Both CD and UC are characterized by exacerbated effector T cell activation [24]. However, both exhibit different T-cell profiles. Whereas Th1/Th17 pathways are predominant in CD, Th2 responses represent the driving force of inflammation in UC [25, 26]. Th1 responses in CD are mediated by IFN-γ which is secreted in response to IL-12 released by antigen presenting cells [27]. This has been recognized initially via experimental colitis models. For example, the use of antibodies against IL-12p40 has alleviated

chemically induced colitis in mice [28]. Furthermore, Hart *et al.* have detected more IL-12-producing dendritic cells (DCs) in the lamina proporia of CD patients relative to control subjects [29]. The pivotal role of Th1 responses in CD was further demonstrated by the improvement of the inflammation in CD and TNBS-induced colitis in mice following the administration of anti-IL-12 and anti-TNFα antibodies [30, 31].

The other inflammatory T helper cell populations involved in the pathogenesis of CD are the Th17 cells [32-34]. The main mediator of Th17 responses in CD is believed to be IL-23 [34]. Initially, genome wide association studies have identified *Il-23R* as a susceptibility gene for CD [35]. Additionally, murine models of either bacterial- or chemical- induced colitis in which IL-23 is deleted or blocked have developed milder colitis as compared to wild type mice [34, 36, 37]. However, the mechanism by which IL-23 modulates the inflammation is debatable. Initially, differentiation of Th17 cells was thought to be induced mainly by transforming growth factor β (TGF-β) and IL-6, while IL-23 acts only to stabilize and/or expand the differentiated cells that generate IL-17 and IL-22 [38]. Nonetheless, IL-23 has showed the capability to generate IL-17 and IL-22-secreting Th17 cells in absence of TGF-β signaling [39]. In addition, a recent study has illustrated that the colitogenic role of IL-23 is independent of IL-17 and IL-22 release [37]. Instead, the IL-23/IL-22 axis showed a protective role against DSS induced colitis via Thy-1 positive innate lymphoid cells [37]. In concordance with these results, IL-17 and IL-22 have shown anti-inflammatory activities against T-cell- and DSS-mediated colitis indicating that IL-23 induces inflammation via a different mechanism [40, 41]. One possibility is that IL-23 promotes T cell dependent inflammation via reducing the differentiation of the anti-inflammatory regulatory T cells (T_{Reg}) [42]. Ivanov *et al.* have illustrated that differentiation of Th17 cells is correlated with

the presence of *Bacteroidetes*, while absence of these bacteria has been accompanied by an enrichment of T_{Reg} cells [43]. Additionally, the transfer of T_{Reg} cells to experimental colitis models has been shown to protect from development of colitis [44]. These observations indicate that gut homeostasis can be described as a balance between Th17 and T_{Reg} cells. Moreover, it illustrates that single taxa of gut microbiota may alter this balance. This anti-inflammatory effect of T_{Reg} cells may arise from the generation of IL-10, IL-35 and TGF-β, or the expression of IL-2 receptor [24].

Regarding UC, the inflammation is thought to be mediated by Th2 cytokines [38]. However, IL-4, the classical marker of Th2 response, is not increased in UC patients indicating that UC has atypical Th2 response [26]. This non-classical Th2 response has been illustrated to be mediated by IL-13 and IL-5 secreted by natural killer T cells instead of conventional T cells [45].

1.3.2. Genetic susceptibility association with IBD.

The role of the genetic factors in the pathogenesis of IBD has been initially recognized via familial aggregation and twin studies [20, 46, 47]. Afterwards, Genome-Wide Association Studies (GWASs) and next generation exome sequencing have identified several risk loci associated with IBD [35, 48-52]. Recently, Jostins and coauthors have increased the number of the identified IBD-associated variants to 163 loci [53]. These loci could be categorized as microbial pattern recognition receptors (e.g. NOD2/CARD15 and TLRs), autophagy (e.g. ATG16L1 and IRGM), Th17 response (e.g. IL-23, CCR6, JAK2 and STAT3), the integrity of the intestinal epithelial barrier (e.g. IBD5, DLG5, PTGER4, ITLN1 and DMBT1), and the unfolded protein response induced by endoplasmic reticulum stress (e.g. XBP1). However, these loci explain a disease variance of only 13.6% and 7.5% for CD and UC, respectively,

indicating that these genes have a minor contribution to the development of the disease [53]. The majority of these identified loci are essential for sensing and responding to gut commensal bacteria and to maintain intestinal homeostasis. Furthermore, these risk factors could affect other antimicrobial innate immune responses such as the production of α-defensin by paneth cells and/or the generation of pro-inflammatory or anti-inflammatory cytokines [54]. Collectively, the identified IBD-susceptibility genes highlight the importance of the interactions between the host immune system and the gut microbiota in the pathogenesis of the disease. However, what triggers this interaction or even how the gut microbes activate the proinflammatory immune responses remains unknown.

1.3.3. Environmental risk factors in IBD.

The second contributors to the pathogenesis of IBD are environmental triggers. The incidence of IBD is higher in developed nations such as Northern Europe, Australia and North America [10, 13-15]. This pattern of distribution may be due to improved diagnosis and/or due to certain environmental factors common in these areas. The extensively studied factors that relate to IBD include diet, smoking, use of antibiotics, stress exposure, sanitary conditions, early gastrointestinal infection, as well as medication and vitamin usage [55-60]. Smoking can be considered the strongest environmental factor correlated with IBD [61]. Smoking is associated with a higher risk of developing CD [57, 62] and with a lower risk of recurrence in ex-smokers compared to people who continue smoking [63]. On the other hand, smoking has been demonstrated as a protective factor in UC [64, 65]. However, how smoking contributes to and/or protects from the disease remains unclear. This contribution may be mediated by modifying the structure of the gut microbiome entirely, as it has been shown recently that smoking affects the structure of the gut microbiota [66]. For instance,

smoking was linked to a decreased relative abundance of *Anaerostipes* which is a known butyrate producing Firmicute [66]. Additionally, side-stream smoking has been illustrated to increase the mouse intestinal *Clostridium* and decrease the gut Firmicutes, *Enterobacteriaceae* and Segmented filamentous bacteria [67]. Another environmental modulator of IBD is diet. It has been revealed that diets with a higher content of fats, poly-unsaturated fatty acids and/or meat increase the risk of CD and UC, while high fiber content, fruits and vegetables are associated with low risk of CD and UC [58]. The association between diet and IBD adds to the role of gut microbiota in IBD as it is known that diet shapes the gut microbiota [68, 69]. For example, rural African children who rely on plant polysaccharide in their diet have lower Firmicutes and higher Bacteroidetes members in comparison with European children [68]. More recently, Martinez *et al.* have illustrated that consumption of whole grain barley or brown rice induces the enrichment of *Roseburia intestinalis, Eubacterium rectale, Roseburia faecis*, and the genera *Dialister*, *Roseburia* and *Bifidobacterium*, with a concurrent decrease of plasma IL-6 [69]. Smith and coauthors have also showed that Malawian children with Kwashiorkor (acute severe malnutrition) exhibit a different microbial profile from their discordant healthy twin [70]. Germ free mice transplanted with the fecal microbiota from these twins and fed on Malawian diet illustrated blooming of *Biolphila wadsworthia* in Kwashiorkor microbiota inoculated mice as compared to healthy microbiota transplanted mice [70]. *B. wadsworthia* is a sulfite reducing microbe that has been linked to IBD [71]. Switching the mice diet to Ready-to-use therapeutic food (RUTF), that includes milk, peanut paste, vegetable oil, sugar, vitamins and minererals, resulted in a prominent increase of beneficial organisms such as *Bifidobacteria, Lactobacilli*, *Ruminococcus* and *Faecalibacterium prausnitzii* [70]. Collectively, these findings indicate that diet shape the configuration of gut microbiota which in turn may affect

the health status of their host. Diet might also have other direct contributions to the development of IBD through antigenic characteristics or by affecting intestinal permeability [72, 73].

The mode of birth delivery has been also known to shape the gut microbiota composition in early life. At the delivery time vaginally born infants showed microbial profiles similar to the vaginal microbiota of their mothers that is dominated by *Sneathia* spp, *Lactobacillus* and *Prevotella*, while C-section delivered infants acquired their mom's skin microbiota that is dominated by *Staphylococcus* spp. [74]. At day 3 of life, the microbiota of caesarean delivery infants are characterized by lower diversity and less *Bifidobacterium* spp. as compared to vaginally delivered infants [75]. At the weaning time, C-section delivered mice exhibited a characteristic microbial profile that is characterized by lower *Rickenellaceae* and *Ruminococcus* and higher abundance of *Bacteroides* and *Lachnospiraceae* as compared to vaginally delivered mice [76]. These disturbances of gut microbiota in early life may affect the immune system maturation which may contribute to the development of IBD. For example, adult mice born through a caesarean procedure showed fewer $Foxp3^+$ regulatory T cells, tolerogenic $CD103^+$ dendritic cells, and less *Il10* gene expression in mesenteric lymph nodes and spleens as compared to vaginally-born mice [76]. However, the association between the mode of delivery and IBD is controversial. While some population-based studies reported caesarean delivery as a risk factor associated with IBD [77, 78], a recent meta-analysis, that involve the results from seven eligible retrospective cohorts and case control studies, reported no association between IBD and the mode of delivery[79].

Another conflicting link has been established between antibiotic usage and IBD. Using population-based data, previous studies have demonstrated an association between antibiotic use, in particular early in life, and the development of both pediatric and adult IBD with odd ratios ranging from 1.32 to 4.94 [80-84]. In contrast, a recent environmental factor questionnaire on Asian and Australian populations illustrated that the antibiotic use is a protective factor for CD and UC development with odd ratios of 0.19 and 0.48, respectively [85]. Treatment with antibiotics has been shown to induce a rapid and significant alteration in the gut microbiota composition such as the loss of microbial diversity [86-88]. Though these perturbations are transient and the microbiota start to stabilize within 4 weeks after ending the treatment [88], the microbiota do not always recover their pre-treatment configurations [86]. Because the gut microbiota shape the immune system during the infancy, early life exposure to antibiotics may induce disturbances to the immune function which may contributes to the development of IBD. In addition to diet and antibiotics [60], vitamins (such as vitamin D) [59] or medications (e.g. steroids and non-steroidal antiinflamatory drugs (NSAID)) [55] have direct contact with gut microbes and intestinal lumen which could lead to either antigenic reactions or inflammatory responses that impact the course of IBD. Other life style factors that contribute to development of IBD and have an impact on the composition of intestinal microbes include: sanitary conditions, urban versus rural living, and family size [56]. **In general, it is clear that the environment and life style have a great impact on the development of IBD. This impact, most likely, results from the alteration of the composition of gut microbiota.**

1.4. Normal gut microbiota.

Our gut is populated by a complex and dynamic microbial community. This community is considered an extra organ within the human body and is collectively known as the gut microbiota. This ensemble consists mainly of bacteria and viruses with a low percentage of archeaeal and eukaryotic members [89, 90]. Using fecal samples from 124 Europeans and via Illumina-based metagenomic sequencing, the gut microbial gene catalogue is estimated to be 150 fold that of human genes, with 99.1% of the genes identified as of being bacterial in origin. The bacterial species that compose the entire cohort were calculated to range from 1000 to 1,150 prevalent species with at least 160 species per individual [89].

Studies of the mucosal and luminal microbiota structure via the sequencing of 16S rRNA clones have revealed that approximately 90% of gut microbes are related to only two phyla: Firmicutes and Bacteroidetes, with the remaining 10% distributed amongst Proteobacteria, Actinobacteria, Fusobacteria and Verrucomicrobia [91]. At a lower level, the majority of the detected Firmicutes are related to the Clostridia class, with the vast majority classified under clostridial clusters IV, XIVa and XVI. Bacteroidetes exhibit less diversity than Firmicutes with 67% of its sequences classified under 3 main phylotypes *B. vulgatus*, Prevotellaceae, and *B. thetaiotamicron*. *B. thetaiotamicron* was the only common *Bacteroidetes* among all tested individuals in that study [91]. Other low abundance species related to other phyla include Proteobacteria species (*Sutterella wadsworthensis*, *Desulfomonas pigra*, *Bilophila wadsworthia*), Actinobacteria genera (e.g. *Actinomyces*, *Collinsella*, and *Bifidobacterium*) and the mucin-degrading Verrucomicrobia species, *Akkermansia muciniphila* [91].

The diversity and the composition of the gut bacteria vary along the length of the gastrointestinal tract (GIT) with a gradual increase in the load and diversity from gastric and duodenum region (10^2 CFU/g content) to colon (10^{11}-10^{12} CFU/g content) [91-94]. The composition also varies at different locations, with *Streptococcus* and *Lactobacillus* species at the duodenum and ending with strict anaerobe such as *Bacteroides*, *Bifidobacterium* and clostridial clusters at the colon [94]. This difference in composition and diversity may be attributed to nutritional substrate availability, oxygen content, acidity and other physiological and immunological conditions at different parts of the GIT.

In addition to the stable differences in the composition observed at different intestinal locations, the gut bacteria are known to, significantly vary among individuals [91]. Moreover, gut bacteria were shown to have continuous temporal variations within the same individual as revealed by the monitoring gut microbiota structure of three subjects for 12 months [95]. This study identified around 40 core species that are persistant over one year in each individual [95]. This core microbiota represent about 75 % of the total community indicating that only small fraction of the gut microbiota is exposed to fluctuations over time [95]. These temporal variations can be attributed to small perturbations in some environmental factors such as short term changes of diet [96]. However, gut microbiota tolerate these temporal changes and return to their original structure once all the physical and the physiological conditions of the gut return to the normal state which is known as "resilience" of the gut microbial ecosystem [96, 97]. Contrary to these temporal changes, the gut microbiota can lose resilience and exhibit permanent compositional changes through exposure to certain factors, such as broad spectrum antibiotics [97], chronic diseases (e.g. diabetes and IBD[93, 98]), obesity[99], or with age [100].

Few aerobes start to colonize the intestine immediately after birth (Enterobacteria, *Staphylococcus* and *Streptococcus*) and these are gradually replaced by anaerobic bacteria untill the microbial profile reaches the same characteristics of the adult gastrointestinal tract [101]. Children from one to seven years of age have a lower fecal microbiota diversity with a higher abundance of Enterobacteria than adults [102]. Additionally, a more recent comprehensive study has indicated substantial differences in microbiota with aging [100], where *Bacteroidetes* is the predominant phyla along with a lower abundance of Firmicutes in the elderly compared with adults. This age-related pattern has been previously confirmed by comparative assessment of the Firmicutes/Bacteroidetes ratio at different ages using quantitative PCR [103]. At lower taxonomic levels, the relative abundance of many bacterial species such as *E. coli*, *Enterococci spp*. *Bacteroides spp. Bifidobacterium spp.* and lactobacilli have been demonstrated to vary with ageing [104, 105]. Recently, a large study has surveyed the gut microbiome structure in the stools of 314 Americans, 114 Malawians, and 100 Amerindians of varying ages [106]. They have reported that subject-based variability is higher in the early stages of life compared with that of the adult, while microbial diversity increases with age [106]. Additionally, they have illustrated that the infant microbiome is dominated by bifidobacteria and that it reaches an adult-like composition by the age of three [106].

1.5. Roles of gut microbiota in human health.

Many reports have described human as a superorganism that lives in symbiosis with different microbes in the various parts of the body [107]. The gut harbors the greatest human microbial ecosystem as it contains trillions of bacterial cells. This ecosystem plays essential roles in human health in general. The first role of the gut microbiota arises from their

metabolic capabilities as they process food constituents including indigestible complex carbohydrates and plant glycans, choline and bile acids [108]. The microbial processing of indigestible polysaccharides generates beneficial short chain fatty acids (SCFA) such as butyrate, acetate and propionate which represent around 90% of the total SCFA produced by the human gut [109]. The major producers of SCFA include *Bacteroides* spp., clostridial clusters IV and XIVa and *Bifidobacterium* spp. Comparative genomic analyses showed that *Bacteroidetes* (*B. thetaiotamicron* and *B. ovatus*) exhibit a high content of glycosidases and lyases in their genome that enable them to utilize indigestable plant and host glycans [110]. Additionally, capillary sequencing of the butyryl-CoA:acetate CoA-transferase gene, which is involved in the main pathway of butyrate production, identified *Eubacterium hallii*, *Eubacterium rectale*, *Roseburia faecis* and *Faecalibacterium prausnitzii* as the most abundant representative butyrate producers in the gut [111]. The highest generation of SCFA in the large bowel starts at the cecum and ascending colon and decreases gradually towards the distal colon [109]. This may be attributed to the accessibility of the substrates which progressively decrease towards the rectum [109]. The majority of the produced SCFA are absorbed by the gut or taken by peripheral tissues such as the liver and muscle [109, 112]. Butyrate provides 10% of the colonocytes energy supply, while acetate and propionate act as metabolic substrates for lipogenesis and gluconeogenesis in peripheral tissues [108, 112, 113]. The contribution of butyrate to colonocytes has been confirmed, as colonocytes of germ free mice have impaired mitochondrial respiration and increased autophagy which can be restored by adding butyrate to the colonocytes [114].

The gut microbiota also contribute to intestinal epithelial integrity in several ways. First, SCFA induce the release of mucin MUC-2 which acts as the first intestinal barrier against invasion of gut microbes [115, 116]. Second, SCFA control gene expression of the

colonocytes either via histone deacetylase (HDAC) inhibition or through the binding to G-protein coupled receptors (GPR41 or GPR43) [108]. For example, sodium butyrate up-regulates the expression of tight junction proteins and their mRNA via the inhibition of HDAC [116, 117]. In addition to SCFA, other microbial components contribute to epithelial integrity. For example, germ free mice have a thinner mucus layer relative to conventionally housed mice which can be corrected following exposure to lipopolysaccharides (LPS) or peptidoglycan [118]. Together, these findings highlight the critical role of the gut microbiota and/or their metabolites in maintaining intestinal barrier integrity.

In addition to the development of tissues and cells, gut microbiota are responsible for the shaping and maturation of the immune system. Germ free mice often have defective gut associated lymphoid tissues (GALT), smaller and fewer Peyer's patches, less cellular lamina propria, fewer cellular mesenteric lymph nodes, lower expression of Toll-like receptors (TLRs) and major histocompatibility complex class II (MHC II) molecules, reduced intraepithelial lymphocytes and $CD4^+$ T cells compared to mice colonized with gut microbiota [119]; Gut colonization of these germ free mice with microbes can correct for some of these deficiencies [119]. Relative to specific pathogen free mice, germ free mice also had an accumulation of invariant natural killer T cells (iNKT) in the lamina proporia of the colon and lung which in turn increased their mortality rate in response to oxazolone-induced colitis [120]. However, the exposure of neonatal, and not adult, germ free mice to commensal bacteria reduced this accumulation and its consequences [120]. Also, mice which received $CD4^+CD62L^+$ lymphocytes from germ free mice developed colitis faster than mice receiving the same T_{Reg} cells from conventionally housed mice [121]. In general, this suggests the critical role of gut microbiota in the development of the intestinal immune system.

The maturation of the immune system in the gut is dependent on the host-specific microbiota. A recent study has illustrated that the colonization of germ free mice with either human or rat microbiota has resulted in fewer intestinal T cells and dendritic cells, and lower antimicrobial peptides expression relative to germ free mice colonized with a murine microbiota [122]. In addition, humanized mice were found to be more susceptible to *Salmonella* infection compared to GF mice colonized with mice microbiota [122]. Both humanized mice and the mice colonized with murine microbiota showed similar bacterial diversity at higher taxonomic levels, but they harbored different bacterial species [122]. This indicates that each host selects a specific microbial consortium that is capable to shape its immune system and maintain the intestinal health.

1.6. Gut microbiota dysbiosis in IBD.

As previously mentioned, our gut is populated with a complex, dynamic ecosystem that contributes to human health in general. Any alteration of the gut microbial consortium may disrupt its functionality and eventually, a diseased state will appear. The dysbiosis of intestinal microbiota is well reported in varying health disorders including irritable bowel syndrome (IBS), obesity, diabetes, and IBD [123-125]. The association between IBD development and gut microbiota dysbiosis was first established through studying animal models of colitis. For example, Garrett and collaborators were able to clearly demonstrate that alteration of gut microbiota composition can induce colitis in genetically intact and immune-sufficient mice [126]. They have reported that a deficiency of T-bet, an important transcriptional factor for gut homeostasis, results in a shift of the microbial population into a colitogenic community. These colitogenic microbiota were able to drive intestinal inflammation in genetically intact mice [126]. In addition to animal models, studies based on

investigating 16S rRNA have shown a decrease in the diversity and dysbiosis of gut microbiota in IBD mucosal specimens compared to healthy biopsies [94, 127]. A single strand confirmation polymorphism (SSCP) fingerprint based on 16S rRNA has shown that microbiota diversity decreased by 50% and 30% in CD and UC, respectively [128]. Frank *et al.* [94] have illustrated this imbalance by sequencing 15,172 small-subunit rRNA (SSU rRNA) clones from 190 biopsies. In that study, Lachnospiraceae and Bacteroidetes were depleted in IBD subjects in association with a relative increase of Proteobacteria, Actinobacteria and bacillus subgroup of Firmicutes. FISH analysis, on the other hand, has illustrated an increase in the relative abundance of Bacteroidetes with a lower abundance of some butyrate producing bacteria such as *Faecalibacterium prausnitzii* in the mucosal biopsies of IBD patients [129-131]. Using 454 pyrosequencing of 16S rRNA V5 and V6 regions extracted from the fecal materials of concordant and discordant twins, ileal CD, colonic CD and healthy individuals have been differentiated from each other according to their respective microbial profiles [132]. Colonic CD has been associated with enrichment of Firmicutes (mainly *Faecalibacterium*, and *Ruminococcus*), Bifidobacteriaceae (*Bifidobacterium*), Coriobacteriaceae (*Collinsella*) and Aneroplasmataceae compared to controls. Conversely, ileal CD showed a depletion of the Ruminococaceae family, especially *Faecalibacterium*, and *Collinsella* with a higher abundance of Proteobacteria (mainly Enterobacteriaceae) relative to healthy microbiota [132]. UC microbiota, on the other hand, have not been differentiated from healthy microbes using the same approach [132]. UC was only associated with depletion of a few species such as *Prevotella*, *Streptococcus*, and *Asteroleplasma* [132]. In contrast, a more recent study has differentiated the microbiome of UC and healthy individuals through sequencing 16S rRNA-clone libraries prepared from the sigmoid colon biopsies of 62 individuals [133]. The microbiota dysbiosis of UC was

characterized by less diversity, fewer Lachnospiraceae and Ruminococcacea families with higher abundance of Proteobacteria and Actinobacteria [133]. **These two studies confirm the association between IBD and gut microbiota dysbiosis. However, the discrepancy between them, especially regarding the UC microbiota structure, may arise from the use of two different sampling approaches**.

Several studies have monitored the gut microbiota composition in pediatric IBD against control children. In 2006, Conte and collaborators [134] reported an alteration in microbiota composition along the length of the intestine in pediatric IBD patients in comparison to control subjects. They followed conventional culturing techniques and 16S rRNA-based real time PCR for quantifying mucosa associated bacteria at the ileum, cecum and rectum of 42 subjects. One important observation in this study was the lower abundance of *Bacteroides vulgatus* in pediatric IBD, the reverse of what was reported for adult IBD [135]. **This result supports the idea that pediatric IBD is a unique type of IBD**. More recently, a study on the relative abundance of 9 bacterial groups using real-time PCR has indicated higher numbers of *E. coli* and lower numbers of *F. prausnitzii* in children with CD compared to control subjects [136]. Furthermore, the microbiota diversity in pediatric IBD patients was shown to differ from that of adults [137]. However, these studies solely examined dominant bacterial groups by applying conventional culture methodology and low-throughput molecular approaches. Since it is known that a high proportion of gut microbiota is uncluturable bacteria [91], it is clear that a comprehensive molecular survey of gut microbiota in pediatric IBD is necessary. Recently, three studies have applied high throughput molecular approaches to characterize the gut microbiota of IBD children [138-140]. The first study used a synthetic learning in microbial ecology (SLiME) analytical approach of 454-pyrosequencing data obtained from fecal samples of 91 individuals and

other published datasets [138]. They were able to discriminate between pediatric patients with IBD and healthy individuals by their gut microbiota composition [138, 140]. More recently, Hansen and coauthors have examined the structure of the mucosal microbiota in 25 children newly diagnosed with IBD and 12 non-IBD subjects using 454-pyrosequencing and real-time PCR [140]. They found that the diversity of the gut microbiota was significantly lower in children with CD but not UC as compared to control subjects [140]. Moreover, they reported that *Faecalibacterium prausnitzii* is enriched in the microbiota from CD subjects as compared to the control group, which is the reverse to what had previously been documented for adult IBD [131]. Recently, Gevers *et al.* examined the bacterial composition and diversity in mucosal biopsies from 447 children with first onset CD and 221 non-IBD subjects [139]. They found that the mucosal microbiota from CD cases were characterized by increased abundance of Veillonellaceae, Enterobacteriaceae, Fusobacteriaceae and Pasteurellacaea, and depletion of Erysipelotrichales, Bacteroidales, and Clostridiales [139]. This microbial dysbiosis was amplified in CD cases treated with antibiotics. For example, the fold changes of Fusobacteriaceae, Enterobacteriaceae, Erysipelotrichaceae, *Bacteroides*, and Clostridiales were increased between CD cases receiving antibiotics and control children [139]. Finally, they illustrated that the microbial profiles identified from mucosal biopsies, but not stool samples, could distinguish CD cases and non-IBD controls [139]. **Overall, these studies support the association between intestinal microbiota imbalance and pediatric IBD. However, a large comprehensive survey of luminal-mucosal microbiota in pediatric UC is still lacking. In addition, the functional link between the microbial dysbiosis and the disease severity is still missing.**

1.7. Different approaches for gut microbiota characterization.

Two main factors should be taken in consideration when studying the gut microbiota: the microbiota sampling approach and the experimental and analytical tools being used**.** Previous studies have relied on either fecal samples and/or intestinal biopsies [129-131, 134, 137, 141]. Fecal samples are widely used because they are easier to collect than biopsies. However, the microbiota characterized from stool differ from the microbiota at the mucosa in terms of both diversity and structure [91, 139, 142]. For instance, several studies [100, 143-146] have shown that Firmicutes outnumber Bacteroidetes by approximately three to four times in stool samples, while the relative abundance between these two major phyla is similar in biopsies [147, 148]. This conclusion was confirmed by studies that compared the microbiota composition of the mucosal and luminal samples from the same subjects [91, 132, 149]. Since we know that the habitat of the colonic microbiota is the loose, outer-mucus layer that can be disturbed or even removed by bowel cleansing procedures prior to biopsies collection [150-152], we can conclude that mucosal biopsies are not always an ideal sampling approach either. Moreover, the majority of the genetic materials extracted from biopsies are of human origin with low microbial content which may affect the quality and accuracy of downstream methodologies and analytical procedures [153]. **These concerns illustrate the need for a new approach to sample the gut microbiota that are in contact with the intestinal barrier and interact with the host immune system.**

Initial investigations of the gut microbiota composition relied on traditional culture techniques such as selective media and phenotypic characterization. However, the unculturable nature of the majority of these intestinal microbes led to incomplete or skewed results [91]. Hence, studies began to switch to culture-independent molecular techniques to

identify and/or quantify the structure of the gut microbiota. Techniques such as terminal-restriction fragment length polymorphism (T-RFLP) and denaturing gradient gel electrophoresis (DGGE) were developed to identify the predominant members of this community [154, 155]. However, these approaches are laborious, time consuming and cannot detect rare microbes. Other quantitative approaches, such as Florescence *In-situ* Hybridization (FISH) [131], real time qPCR analysis [134, 136], and sequencing of 16S rRNA clones libraries [91, 94], have improved our ability to quantify dominant members and/or certain constituents of mixed populations, but are still not suitable for the comprehensive analysis of complicated communities. Moreover, these approaches are limited by the number of samples that can be analyzed simultaneously. Recently, high throughput analysis tools such as microarray and next generation sequencing of highly variable regions of the 16S rRNA gene have been developed to characterize the gut microbiota [144, 156-159]. These approaches allow for the detection of hundreds of thousands of DNA sequences, while simultaneously enabling comprehensive analysis of gut microbes on larger scales.

1.7.1. 16*S* rRNA-based microarray.

DNA microarray approaches can simulatenously detect thousands of individual DNA sequences. Hence, it has been recognized as a valuable method for high throughput screening and quantitation of microbiota [160, 161]. The application of microarray technology in characterizing the gut microbiota has improved dramatically over the past decade. In the earlier days, only 60 16S rRNA-based probes were designed to identify 20 bacterial genera in human feces [162]; then the number of the targeted species increased to 40 species by using three 40-mer probes per species [161]. The first comprehensive 16S- based microarray

(10,462 rDNA probes) was developed by Palmer *et al* in 2006 [160]. This microarray was later used to monitor the development of the infant microbiota during the first year of life [101]. Each of these early microarrays relied on long 40 mers oligonucleotide probes. More recently, shorter (16-21 mers) oligonucleotide-based microarrays were developed, which exhibited higher sensitivity as compared to the previous microarrays based on 40-50 mers probes [159]. The use of 16S rRNA-dependent microarrays was the method of choice at the early stages of studying the gut microbiota as they were able to produce a large amount of information over a short period of time. Additionally, they were considered relatively low cost and their data analytical tools were initially more mature in comparison with deep sequencing. However, because microarray probes are designed from known genomic data, they could not identify novel bacterial species for which 16S rRNA sequence is not available.

1.7.2. 16S rRNA sequencing approaches.

Current gut microbiota studies mainly rely on 16S rRNA-based sequencing approaches. These approaches exhibit direct access to the sequences and therefore are able to detect novel microbes. Initially, amplicon cloning followed by capillary sequencing of clone libraries was employed to characterize the gut microbiota structure [91, 163]. This approach had scaled up from the sequencing of 100 clones in 1996 [163] to 15,172 clones in 2007 [94]. In spite of the merit of identifying novel bacteria (for example Eckburg and coauthors detected 244 novel bacteria out of 395 total identified bacteria [91]) and the high confidence and resolution achieved by analyzing the entire sequence of 16S rRNA gene, this approach is limited by its high cost, the number of samples and/or the number of clones that could be analyzed, in addition to being very time intensive. Now, the revolution of next generation

sequencing approaches has eased these obstacles and has enabled the comprehensive characterization of the gut microbiota. These approaches depend on the extraction of metagenomic DNA and the amplification and sequencing of one or more hypervariable sequence from the 16S rRNA gene. The diversity and the relative abundance of the amplified sequence variants are then used to describe the diversity and the composition of the community. The 16S rRNA gene consists of nine hypervariable regions (V1-V9; Figure 1.1) that exhibit sequence variability among different bacterial species [164]. These regions are flanked by conserved sequences which enable the PCR amplification of one or more hypervariable region using universal primers [165]. The two main sequencing technologies used are the 454-pyrosequencing and the Illumina sequencing. Although full length sequencing provides higher resolution and confidence for taxonomic assignment, Next generation sequencing reduces the cloning bias associated with cloning and capillary sequencing approaches. The use of the Illumina GAIIx platform has increased the number of sequences generated per run to $1x10^8$ compared to the $5x\ 10^5$ reads generated by Roche 454 GS FLX Pyrosequencing and $1x10^3 – 1x10^4$ sequences from Sanger sequencing [166]. Now, the Illumina Hiseq 2500 platform is capable to produce a maximum of 6 billion paired end reads per run (Illumina Inc., 2013). This sequencing depth allows for multiplexing of samples and the identification of very rare phylotypes. In addition to 16S rRNA-based phylotypes detection, Illumina sequencing has been applied to metagenomic studies of the gut microbiome. For example, Qin and coworkers identified 3.3 million microbial genes by analyzing 576.7 gigabases obtained via Illumina GA sequencing of 3 metagenomic libraries derived from the fecal samples of 124 Europeans [89]. This study shows the depth of sequencing achieved by Illumina, which compensates for the short read length generated by this approach [89]. To increase read length, new protocols such as paired end Illumina

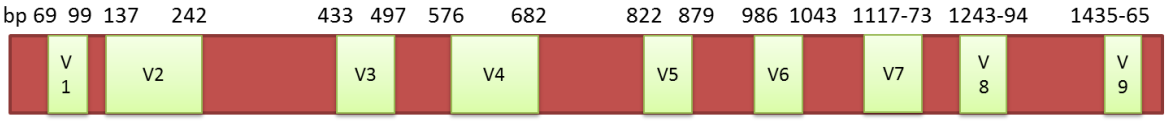


Figure 1.1: Schematic representation of the conserved and hypervariable regions of 16S rRNA gene. The conserved regions are shown in brick red while the hypervariable regions (V1-V9) are shown in light green. The start and ending locations of each variable region is illustrated on the top of them as determined by Chakravorty *et al*. [165].

sequencing have been introduced, which involves sequencing of both ends of a fragment, thus increasing the read length up to 200 bases per read after merging [167]. In general, these advances in next generation sequencing approaches have lowered the cost of sequencing (<11 USD per sample), increased the sensitivity for detection of rare phylotypes and minute temporal variations among samples and enabled the comprehensive analysis of large numbers of samples [166].

1.8. Dysfunctionality of gut microbiota in IBD.

The gut is populated by a complex ecosystem which exhibits many beneficial functions to the host including nutritional, developmental and immunological roles that collectively contribute to human health**.** The developments of high throughput sequencing tools have revealed the association between dysbiosis of this ecosystem and IBD [93, 94, 130, 140, 168, 169]. Whether this dysbiosis is a result or a cause of inflammation remains to be elucidated. Also, the consequences of this dysbiosis on the chronicity of inflammation are unknown. It is difficult to functionally associate one bacterial species to inflammation especially if this species colonizes the gut at a low abundance. Additionally, bacteria share common functions, and thus a dysbiosis of certain bacterial taxon does not mean a significant change in the biota functionality. Hence, in addition to investigate the variation in microbial composition, it is also necessary to assess the functional association between the gut microbiome and IBD. That is why the trend of linking the gut microbiome to IBD has begun to switch from 16S rRNA-based phylogenetic characterization to metagenomic analysis of the microbiome either at the community or species level. Recently, two metagenomic studies have attempted to link the dysfunctionality of the gut microbiome to IBD using two different approaches [66, 124]. First, Greenblum *et al.* [124] analyzed shotgun metagenomic data from

124 individuals and mapped the metabolic network of the entire gut community. Next, they calculated the odds ratios and compared the topology of the metabolic pathways, genes and biological processes of the gut microbiota from obese and IBD subjects to healthy individuals [124]. This approach revealed an association between IBD and the genes involved in NO_2 production as well as the metabolism of choline and p-cresol [124]. Secondly, Morgan *et al.* first identified the microbiota composition of each disease phenotype using 16S rRNA-based sequencing, then they inferred the metagenome of each microbial community by constructing gene catalogue for each phenotype, and finally they assessed the identified genes and pathways through sparse multivariate analysis [66]. They further confirmed their findings by shotgun metagenomic analysis of the fecal microbiota from 4 CD and 7 healthy subjects [66]. This study reported an IBD-associated enrichment of genes involved in glutathione transport and metabolism, sulfur amino acid metabolism, redox homeostasis, mucin degradation, secretion systems, adhesion and invasion as well as a depletion of genes involved in biosynthesis of some nucleotides, amino acids and SCFA [66]. **Collectively, the microbial dysbiosis associated with IBD will result in impaired metabolic activity of the gut microbiota. Identification of the gut microbiota dysfunctionality in first onset pediatric IBD remains unclear.**

1.8.1. Hydrogen sulfide metabolism in IBD.

The gut microbiota generate hydrogen sulfide (H_2S) via 2 different biochemical pathways. Firstly, sulfate reducing bacteria (SRB) couple the reduction of sulfate as a terminal electron acceptor to the oxidation of H_2 and/or other organic compounds (e.g. lactate) as an electron donor during their anaerobic respiration [170]. The second pathway is employed by bacteria that are able to degrade sulfur containing amino acids. These bacteria

depend on desulfhydrases such as cysteine desulfhydrase or other enzymes able to metabolize sulfur containing amino acids [171]. Sulfate reducing bacteria include members of *Deltaproteobacteria* along with the genera *Desulfotomaculum, Desulfosporosinus*, *Thermodesulfobacterium*, and *Thermodesulfovibrio* [172]. Using multiplex PCR to identify fecal SRB isolates, Loubinoux *et al.* have reported that luminal SRB are dominated by *Desulfovibrio piger*, *D. fairfieldensis*, and *D. desulfuricans* [173]. More recently, a diversity analysis of *Desulfovibrio* spp using the dissimilatory sulphite reductase (*dsrAB*) gene confirmed the predominance of *D. piger* with the detection of a new, unclassified SRB [174]. In another study, 454 pyrosequencing of *dsrAB* gene fragments identified 8 more sulphate- and sulphite-reducing bacteria in addition to the sulphite reducing bacterium, *Bilophila wadsworthia* [175]. The identified bacteria included four known species; *D. piger*, *Desulfovibrio sp.* NY682, D. *vulgaris*, and *D. desulfuricans F28-1*; and four new species that are highly similar to *D. desulfuricans* F28-1 (93% *dsrAB* sequence similarity); *D. oxamicus* (84% identity); *Desulfotomaculum sp. Lac2* (80% identity); and *D. simplex* (88% identity) [175]. On the other hand, the gut microbiota that generate H_2S through the fermentation of aminoacids include *Fusobacterium nucleatum*, *Atopobium spp.*, *Gemella sanguinis*, *Micromonas micros*, *Streptococcus spp.*, *Actinomyces spp*., *Eubacterium spp.*, *Veillonella spp.*, *Bulleidia moorei.*, *Prevotella spp.*, *Campylobacter spp.*, and *Selenomonas spp*. [176]. In addition to the hydrogen sulfide released by the microbiota, H_2S can be biosynthesized endogenously by the intestinal colonocytes from L-cysteine via two enzymes: cystathionine β-synthase and cystathionine γ-lyase [177]. Two additional pathways (3-mercaptopyruvate sulfur-transferase in combination with cysteine amino-transferase and 3-mercaptopyruvate sulfur-transferase coupled with D-amino acid oxidase) have been recently identified as associated with H_2S production peripherally [177]. The normal concentration of luminal H_2S

ranges from 1.0 to 2.4 mmol/L in the human large intestine [178], while the concentration in fecal contents ranges from 0.17 to 3.38 mmol/kg [179, 180]. Taking into consideration the lipid solubility and the passive diffusion of H_2S through the intestinal mucosa, these concentrations are likely underestimated.

At low concentrations (< 1mM), hydrogen sulfide is considered a cytoprotective metabolite that induces some cellular anti-inflammatory responses [181]. These cellular responses include, for example, prevention of caspase activation and apoptotic cell death [182], inhibition of leukocyte adhesion to vascular endothelium which in turn decreases the infiltration of neutrophils and lymphocytes [183], induction of cyclooxygenase-2 (COX-2) expression [184], and promotion of neutrophil apoptosis [185]. The last claim has been challenged by Rinaldi and coworkers who demonstrated that H_2S induces a short term survival of neutrophils which may potentially accelerate the resolution of the inflammation process [186]. Wallace and coauthors illustrated that colonic biosynthesis of H_2S is significantly increased after induction of colitis [184]. Inhibition of H_2S synthesis was shown to exacerbate TNBS-induced colitis and triggers colonic inflammation in rats [184]. Additionally, treatment of rats with H_2S donors (such as NaHS) promotes repair of damaged tissues and accelerates resolution of experimental inflammation [184, 187]. In another *in-vitro* study, H_2S at normal colonic concentrations lowered the proliferation of cancerous and normal colonic cells and induced autophagy through the AMP-activated protein kinase (AMPK) pathway [188]. In contrast, higher concentrations of H_2S act as genotoxic and/or cytotoxic transmitter by affecting genes that are responsible for cell cycle progression, DNA repairs and inflammatory responses [189]. The main cytotoxic effect of H_2S is the inhibition of cytochrome c oxidase activity which is the terminal oxidase of mitochondrial respiration [190, 191]. These effects prevent the oxidation of essential metabolites such as n-butyrate, L-

glutamine and acetate which eventually decrease the bioenergetic performance of the cell [190, 191]. It has been shown that sulphide inhibits butyrate oxidation in rat colonocytes through the inhibition of short chain acyl dehydrogenation of activated fatty acids [192]. This antagonistic effect of H_2S on butyrate oxidation has been shown to induce hyperproliferation of the colonic mucosa [193]. A recent study has reported an indirect cytotoxicity of hydrogen sulfide, where increased sulfide production induces the conversion of nitrite to NO which has a damaging effect on the colonocytes [194].

To keep the local concentration of H_2S at a harmless level, the colonic mucosa expresses a specialized H_2S oxidation system that degrades H_2S to sulphate and thiosulphate [195]. This oxidation system consists of the enzymes sulfide quinone reductase (SQR), ethylmalonic encephalopathy (ETHE1), and thiosulfate sulfur transferase (TST which also is known as rhodanese) [196]. The mitochondrial oxidation of H_2S is not essential for cellular respiration, but instead, is required for the detoxification of the excess H_2S [196]. In contrast, the respiratory capacity of the cell is an important parameter that affects the efficiency of H_2S detoxification [196].

A higher concentration of H_2S has been proposed to be associated with the pathogenesis of IBD [71, 197]. The higher concentrations could arise either from an increased abundance of H_2S-producing bacteria or from a deficiency in the H_2S-detoxification pathway. The association between H_2S and IBD was first reported for UC [198]. In 1998, Levine and coauthors demonstrated that the release of H_2S from fecal samples was 3-4 fold higher in UC cases as compared to control subjects [198]. Furthermore, incubation of rectum biopsies from UC patients with sodium hydrogen sulfide resulted in increased proliferation of colonic mucosa, while hyperproliferation was reversed when the

biopsies were co-incubated with sulfide and butyrate [193]. These results reveal a toxic role of H_2S on the colonic mucosa; this toxic role, most likely, arises from disturbing butyrate metabolism. Treatment of UC patients with 5- aminosalicylic acid-containing drugs lowered the production of hydrogen sulfide as indicated by fecal sulfide level; this consequencce was proposed to contribute to the therapeutic activities of these drugs [199]. Moreover, sulfate reducing bacteria were isolated from 80% of UC pouches but not from patients with familial adenomatous polyposis [200]. One of these bacteria; *Bilophila wadsworthia*, has been recently shown to potentiate the development of colitis in *Il-10*$^{-/-}$ mice [71]. However, this association between hydrogen sulfide and UC remains controversial as several studies could not identify any significant difference in fecal or mucosal sulfide and/or SRB abundance between control and UC subjects [201-203].

The association between hydrogen sulfide and UC is also supported by indirect evidence. First, a high protein intake, which results in elevated microbial H_2S production, is known to be associated with a higher risk of IBD [204]. Second, UC patients are characterized by a higher activity of fecal mucin sulphatase as compared to controls [205]. Mucin sulphatase releases sulphate from mucosal sulfomucin and this endogenous sulphate may provide a source for H_2S biosynthesis by SRB [205]. On the other hand, the link between CD and hydrogen sulfide is still missing. Jia and coworkers have reported no difference in the general abundance of SRB between CD and healthy subjects [175]. However, some indirect evidence also postulates the association between H_2S and CD. For example, the increased metabolism of sulfur containing amino acids, such as methionine and cysteine, along with a decrease in the metabolism of other amino acids, such as lysine and glutamine, has been associated with ileal CD [66]. Also, the same study reported an over-

representation of microbial sulfate transport genes in CD patients. Both sulfate and sulfur containing amino acids are the main precursor for hydrogen sulfide biosynthesis as previously mentioned.

The second factor that could contribute to a higher colonic sulfide concentration is the dysfunctionality of H_2S detoxification. The dysfunctionality of H_2S detoxification genes in IBD remains debatable. First, Pitcher and coauthors have shown that the activity of thiol-methyl transferase (TMT) is higher in the peripheral blood of UC patients [206]. This was supported in 2007 by the study of Picton *et al*., which demonstrated that the activities of TMT and rhodanese were higher in the erythrocytes but not in the rectal biopsies of UC patients relative to healthy volunteers [207]. In contrast, no differences were detected in the activities of these enzymes in the erythrocytes or the rectal biopsies of CD patients as compared to controls [207]. In 2009, a potential role for H_2S detoxification in IBD resurfaced when the activity and the expression of rhodanese was shown to decrease in parallel to the development of inflammation in the dextran sodium sulfate-induced colitis mice model [208]. In concordance with previously mentioned studies, this loss of activity in the colon was followed by an increase in its activity in red blood cells [208]. Additionally, impaired detoxification of hydrogen sulfide was also observed in UC patients through the assessment of the expression level and activity of TST in colonic mucosal biopsies [209]. In general, the association between hydrogen sulfide and IBD remains controversial, either at the gas generation or detoxification levels.

1.8.2. Butyrate metabolism in IBD.

Butyrate is the conjugate base of butanoic (butyric) acid, a weak acid with a pKa of 4.8. Considering the neutral pH of the intestine, the majority of butyrate will be in the

anionic form rather than the free acid form. Butyrate biosynthesis by the gut microbiota starts by the condensation of two molecules of acetyl CoA to generate one molecule of butyryl CoA. Next, butyrate is generated from butyryl CoA via two main possible pathways: the butyrate kinase and phosphotransbutyrylase or butyryl-CoA: acetyl-CoA- transferase [210]. The second pathway has been illustrated to be predominant among butyrate producing microbes in the human colon [211]. The major butyrate producing bacteria that inhabit the human gut are related to clostridium clusters XIVa and IV with a few percentage of clostridium clusters I, XV and XVI [210]. Using a continuous anaerobic culture of fecal microbiota at variable pH and peptide supply, it has been shown that the majority of butyrate is generated at pH 5.5 and is decreased significantly when the pH is increased to 6.5 [212]. In parallel, the proportions of the major butyrate-producers, *F. prausnitzii* and *Rosuberia*/*Eubacterium* spp, were markedly increased at pH 5.5 compared to pH 6.5 [212]. The intraluminal pH of the human gastrointestinal tract increases gradually from 5.7 in the cecum to 6.7 in the rectum [213]. This indicates that the cecum/proximal colon area is the main location for butyrate production by butyrate producing bacteria.

The association between IBD and impaired butyrate metabolism has been reported in many prevous studies [209, 214-217]. First, the oxidation of butyrate was shown to be significantly impaired in UC colonic tissues as compared to healthy controls [217]. As a result of impaired butyrate oxidation in UC subjects, the luminal butyrate concentration was significantly increased in ulcerative colitis as compared to controls and this raised concentration was strongly correlated with the severity of mucosal inflammation [217, 218]. Additionally, a depleted butyrate uptake by colonocytes in UC subjects has been shown to result from the lowered expression of the butyrate transporter: monocarboxylic acid

transporter 1 (MCT-1) [214, 219]. The impaired butyrate uptake and β-oxidation in UC have been confirmed through the assessment of gene expression levels in colonic mucosal biopsies of 88 UC subjects [214]. The same group linked the dysfunctionality of butyrate oxidation to impaired hydrogen sulfide detoxification which then leads to a high cellular concentration of sulfide [209]. In parallel, butyrate intake has shown a protective effect against the development of colitis that is independent of restoring butyrate oxidation [215]. For example, administration of butyrate either orally or via local enema to mice or rats has alleviated the chemically induced mucosal inflammation [220, 221]. Butyrate may exhibit this anti-inflammatory role through various mechanisms. First, butyrate may inhibit NFκB activation, which results in the suppression of proinflammatory cytokines [222, 223]. This inhibitory action is impaired in IBD individuals as illustrated by the assessment of butyrate effect on cytokines production by peripheral blood mononuclear cells (PBMC) isolated from IBD and healthy subjects in response to TLR-2 activation [216]. Secondly, butyrate may induce Fas-mediated apoptosis of T cells via inhibition of HDAC-1 [224]. This in turn may inhibit IFN-γ-induced STAT1 activation which would result in reduced colonocyte expression of inflammatory mediators such as nitric oxide synthase (iNOS) and cyclooxygenase 2 (COX2) [224]. Alternatively, butyrate may modulate the antioxidant machinery of the cell by stimulating the expression of the glutathione-S-transferase-alpha and inhibiting reactive oxygen species (ROS)-mediated activation of NFκB [225]. Anti-oxidative effects of butyrate also include up-regulation of human catalase and metallothionein and depression of cyclooxygenase 2 and glutathione reductase expression [226]. Finally, butyrate is known to reduce inflammation by contributing to intestinal barrier integrity. For example, intra-rectal inoculation of *Clostridium tyrobutyricum*, a potent

butyrate producer, restored MUC-2 secretion, expression of zonula occludens (ZO)-1 tight junction protein and reduced the release of cytokines in DSS treated mice [116].

The role of butyrate in IBD has also been indicated by the assessment of the gut microbiota composition. For example, screening of fecal microbiota of 6 healthy and 6 CD individuals via 16S rRNA clone sequencing and FISH analysis has revealed a depletion of the *Clostridium leptum* cluster in CD cases compared to a healthy group [227]. In agreement with this, Sokol and coauthors have illustrated that *F. prausnitzii* was depleted in CD mucosal microbiota and this reduced abundance was associated with a higher risk of recurrence of ileal CD [228]. Although, their study also revealed that the anti-inflammatory activity of this bacterium was independent of butyrate production. Recent studies have confirmed the depletion of *F. prausnitzii* and other butyrate producers in the fecal microbiota of both CD and UC subjects compared to their unaffected relatives and other healthy controls [168, 229]. In contrast to adult populations, the depletion of *F. prausnitzii* in pediatric IBD is still controversial. While Hansen *et al*. showed an increased relative abundance of *F. prausnitzii* in pediatric CD patients as compared to healthy subjects [147], Gevers and coauthors recently showed that *F. prausnitzii* is depleted in CD children [139]. These two studies characterized the gut microbiota composition in pediatric IBD at the time of diagnosis. However, they sampled the mucosal microbiota from different intestinal regions; Hansen *et al*. relied on distal colonic biopsies, while the study of Gevers *et al*. employed terminal ileal and rectal biopsies [139, 140]. These findings suggest that the biogeographic features of the gut microbiota are an essential parameter to note when interpreting such data.

1.9. Hypothesis and objectives of this study.

An intricate and essential partnership is established early in life between the host and the intestinal microbiome, assuring the maintenance of microbiota homeostasis. Disturbance of this relationship is often associated with various pathological conditions including inflammatory bowel diseases (IBD). Early onset IBD constitutes a unique phenotype of the disease with specific clinical outcomes, epidemiological distribution and genetic background. Although the association between IBD in adults and the gut microbiota dysbiosis is well established, a comprehensive characterization of this ecosystem in first onset CD and UC at childhood is still lacking. Moreover, our understanding regarding the contribution of this microbial ecosystem to disease severity and the cause/effect relationship between the disease and intestinal microbiota has so far been limited.

The first objective of the study described herein is to compare the microbial profile generated from mucosal aspiarates with that of mucosal biopsies and stool samples. Secondly, we performed a comprehensive high throughput characterization of the gut microbiota diversity and structure in a first-onset cohort of pediatric IBD patients using mucosal aspirates as microbiota sampling method. Finally, we aimed to gain a better understanding of the host-microbe metabolic interactions that may contribute to the inflammation via murine colitis models and Super-SILAC analysis of the host metaproteome. I hypothesized that the gut microbiota composition is a characteristic of each IBD subtype (CD and UC) at the time of diagnosis, and differs from that of healthy individuals. I equally hypothesized that alteration of the gut microbiota structure in IBD contributes, metabolically, to disease severity.

CHAPTER 2. Materials and Methods

2.1. Ethics statement.

The children protocol was approved by the Research Ethics Board of the Children's Hospital of Eastern Ontario (CHEO). The protocol for the use of germ free (GF) mice was approved by the Animal Care Committee of the University of Ottawa. The protocol for the use of *Il-10*$^{-/-}$ *;NF-kB*^{EGFP} mice was approved by the Institutional Animal Care and Use Committee of University of North Carolina at Chapel Hill.

2.2. Subject selection and sampling.

All patients under 18 years of age and scheduled to undergo their first diagnostic colonoscopy at the Children's Hospital of Eastern Ontario (CHEO) were considered potentially eligible for recruitment into this study, with the following exclusions which are known to affect the gut microbiota composition: (1) a body mass index (BMI) greater than the 95th percentile for age; (2) diagnosis with diabetes mellitus; (3) diagnosis with infectious gastroenteritis within the preceding 2 months; and (4) the use of any antibiotics, probiotics or immunonodulatory medications within the last 4 weeks. The same exclusion criteria were applied to the non-IBD control group. All IBD cases were newly diagnosed with IBD (inception cohort prior to the initiation of treatment) and met the standard diagnostic criteria for either ulcerative colitis or Crohn's disease following thorough clinical, microbiologic, endoscopic, histologic and radiologic evaluation [230]. Phenotyping of disease was based on endoscopy and clinical disease activity scores. The Simple Endoscopic Score for Crohn's Disease was used to record macroscopic activity in each segment of the intestinal tract in Crohn's disease [231], the site of involvement in CD was recorded utilizing the Paris Modification of the Montreal Classification for IBD [232] and clinical disease activity of CD was determined using the Pediatric Crohn's Disease Activity Index (PCDAI) [233]. For UC,

the disease site was recorded using the Paris Modification of the Montreal Classification for IBD [232], endoscopic activity was recorded using the Mayo Score Flexible Proctosigmoidoscopy Assessment in ulcerative colitis and clinical activity of UC was determined using the Pediatric Ulcerative Colitis Activity Index (PUCAI) [234]. All controls had a macroscopically and microscopically normal colon, and did not carry a diagnosis for any known inflammatory intestinal disorder or any condition that is known to be associated with gut microbiota dysbiosis such as irritable bowel syndrome (IBS), celiac disease and infectious diarrhea. Data collected on all participants included: demographics (age, gender, BMI, country of birth, age of diagnosis), environmental exposures (cigarette smoke, diet, previous antibiotic exposure), and all clinical features.

Mucosal-luminal interface samples were collected from the left colon (LC; n= 39, 34 and 13 for CD, control and UC, respectively), right colon (RC; n= 93, 60 and 36 for CD, control and UC, respectively) and terminal ileum (TI; n=27, 15 and 11 for CD control and UC, respectively) at the time of endoscopy. Colonoscopy preparation was done the day before the procedure as per standard protocol [235]. During endoscopy, once the correct position is reached, loose fluid and debris was aspirated. Thereafter, sterile water was flushed onto the mucosa and the collection of water, mucus and intestinal cells of the colonic mucosa was aspirated into a sterile container through the colonoscope. These samples were immediately placed on ice in the endoscopy suite, promptly transferred to the lab to minimize delay for processing and then stored at -80°C. Up to 2 biopsies were collected from the macroscopically involved and non-involved areas of the right colon. Biopsies were flash frozen on dry-ice in the endoscopy suite and immediately stored at -80°C until further processing. Stool samples (n=5) were collected by the patients one day before the

colonoscopy and immediately frozen. The stool samples were brought to the CHEO the next day and immediately stored at -80^{o}C until further processing.

2.3. Metagenomic DNA extraction.

Metagenomic DNA was extracted using the Fast DNA Spin Kit (MP Biomedicals, Solon, OH) and the FastPrep machine (MP Biomedicals, Solon, OH) with two mechanical lysis cycles at speed 6.0 for 40 seconds, with 5 minutes cooling on ice between the two cycles. Next, the DNA was isolated and purified as per Fast DNA Spin Kit protocol. The extracted DNA was quantified by QuBit fluorometer (Invitrogen, Carlsbad, CA) and stored at -20ºC until use.

2.4. 16S rRNA-V6 454-pyrosequencing.

Preparation of the 16S rRNA-V6 library for 454 pyrosequencing and generation of the sequences was conducted in collaboration with Dr. Mehrdad Hajibabaei and Dr. Shadi Shokralla at the Biodiversity Institute of Ontario, Department of Integrative Biology, University of Guelph, following their standard protocol. Briefly, the sequencing library of the 16S rRNA-V6 region was constructed using a two-step PCR strategy using the primers; 16SF 5'-AAACTCAAAKGAATTGACGG-3' and 16SR 5'-ACGAGCTGACGACARCCATG-3'. First, ten replicates from each sample were amplified in a 25 µl reaction containing 50 µg DNA template, 1x PCR buffer, 2 mM $MgCl_2$, 0.2 mM dNTPs mix, 0.2 mM of each primer and 2.5 U Platinum Taq polymerase (Invitrogen, Carlsbad, CA). The reactions were heated to 95°C for 5 min, followed by 15 cycles of 94°C for 40 sec, 46°C for 1 min and 72°C for 30 sec, and a final extension at 72°C for 5 min. The amplicons from each sample were pooled and purified using QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) following its standard protocol. Two µl of the pooled amplicons were used as a template for the second PCR following the same conditions except by using

30 cycles of amplification instead of 15 cycles. The second PCR aimed to add the titanium tails to the V6 amplicons which are essential for the 454 sequencing procedure. The amplicons from the second PCR were pooled and purified again using the QIAquick PCR Purification Kit. The purified amplicons from different samples were normalized to the same concentration (100ng/µl), captured to streptavidin coated sepharose beads and exposed to emulsion PCR. Afterwards, the immobilized double stranded amplicons were denatured into single stranded DNA which was then annealed to the sequencing primer by heating at 65°C for 5 min. The single stranded DNA-containing beads were exposed to pyrosequencing on a 454 Genome Sequencer FLX System (Roche Diagnostics GmbH) using GS Titanium chemistry according to the standard amplicon sequencing protocol. Twelve samples covering the three tested phenotypes (4 samples each) were each sequenced on the same run with each being sequenced in a 1/16 section of 70 × 75 picotiter plate.

2.5. 454-pyrosequencing data analysis.

A total of 346,160 reads were generated from the 454 pyrosequencing of 16S rRNA-V6 region from 26 right colon samples. The raw pyrosequencing reads were processed to remove low quality and short reads using Quantitative Insights Into Microbial Ecology pipeline release 1.4.0 (QIIME 1.4.0) [236] according to the following parameters: (1) minimum read length of 100 bp, (2) exact matching to the sequencing primers, (3) no ambiguous nucleotides, and (4) a minimum average quality score of 20. Next, sequences were clustered using UCLUST [237] based on average percentage of identity of 97%. The most abundant read from each operational taxonomic unit (OUT) was picked as a representative sequence for that cluster. We then used Pynast to align the representative sequences with a minimum alignment length of 100 and a minimum percentage identity of 75%, followed by checking the chimeric OTUs with the blast_fragments approach. Only 6

representative sequences were identified as chimeric and therefore were removed. Taxonomy assignments were made with BLAST [238] by searching the representative sequences against the Greengenes database (release 4Feb, 2011) with an e value of 1e-8 and a confidence score of ≥ 0.5. Next, singletons (OTUs that had only one matching sequence) were filtered out from the resulting OTU table. The OTU table was then used to determine the alpha and beta diversity within and between the samples using the default criteria of QIIME. To identify the microbial biomarkers for each tested category, the relative abundance of different phylogenetic levels computed by QIIME was analyzed by Linear Discriminant Analysis (LDA) Effect Size (LEfSe) algorithm (http://huttenhower.org/galaxy/) [239] using its default parameters. The 454- pyrosequencing reads assigned as *Atopobium* by QIIME analysis were retrieved and found to match to *A. parvulum* following alignment of the reads against the RDB [240] and NCBI databases (the aligned region covered the entire 454 sequence length with >99% sequence identity to *A. parvulum*).

2.6. 16S rRNA-V6 library construction for Illumina sequencing.

A two-step PCR strategy was used to amplify the V6 hypervariable region of the 16S rRNA gene. Briefly, the V6 region of 16S rRNA was amplified using two successive PCR reactions. The first PCR added the Illumina paired-end sequencing adapters and the barcode sequences using modified universal 16S rRNA-V6 primers [241] (Appendix I). Each reaction was prepared in a total volume of 50 µl using 50 ng of the extracted DNA, 0.5 µM of each primer, and 1x Phusion Flash High-Fidelity PCR Master Mix (Thermo Scientific, Vilnius, Lithuania). The reaction was heated to 98°C for 30 sec and then subjected to 10 cycles of 98°C for 5 sec, 61°C for 15 sec with 1°C drop each cycle and 72°C for 15 sec, followed by additional 15 cycles with an annealing temperature of 51°C for 15 sec, and a final extension at 72°C for 5 min. The second PCR was carried out using 10 µl of the first

PCR products in a final volume of 50 µl using the primers PCRFWD1/PCRRVS1 (Table 2.1). The PCRFWD1/PCRRVS1 primers are complementary to the flow cell primers at the 5' end and Illumina paired-end sequencing adapter at their 3' end. The second PCR conditions were 1 min at 98°C, 15 cycles of 10 sec at 98°C, 30 sec at 65°C, and 30 sec at 72°C followed by a final extension step at 72°C for 5 min. The amplicons of each sample were visualized on 1.5% agarose gel and purified using the Montage PCR_{96} Cleanup Kit (Millipore, Billerica, MA). Afterwards, the DNA concentration in each reaction was quantified using the Qubit® dsDNA BR Assay Kit (Invitrogen, Carlsbad, CA) according to the manufacturer instructions. Finally, equimolar quantities of the amplicon from all samples were pooled, gel purified using QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) and the libraries were sent to The Center for Applied Genomics (TCAG), Toronto, for paired-end Illumina HiSeq2000 sequencing.

2.7. Illumina sequencing data analysis.

Paired-end sequences obtained by Illumina HiSeq 2000 were merged into longer reads using Fast Length Adjustment of Short reads (FLASh) software, while avoiding any mismatch in the overlap region that ranged from 10 to 80 nucleotides [242]. More than 95% of the reads were successfully merged, while the sequences that failed to merge were discarded. The merged reads were then quality filtered with a minimum quality score of 20 using the fastq_quality_filter command from the Fastx toolkit (http://hannonlab.cshl.edu/). High quality reads were sorted according to the forward and the reverse barcode sequences with barcode trimming using the NovoBarCode software (Novocraft.com). Next, the reads were fed to QIIME 1.7.0 [236] to determine the taxonomic and diversity profiles of the samples. First, reads were clustered into OTUs using a closed-reference OTU picking workflow against the Greengenes reference set (release 4Feb, 2011) based on an average

percentage of identity of 97%. The resulting OTU table was used to summarize the taxonomy and to assess the alpha and beta diversity within and among the samples using the default criteria of QIIME. To estimate the microbial diversity captured from each sample, we calculated the Good's Coverage score by using all the identified OTUs. To determine the microbial biomarkers of each class, the relative abundance of different taxa computed by QIIME was analyzed using the LDA Effect Size (LEfSe) algorithm (http:// huttenhower.org/galaxy/) [239]. The phylogenetic trees of core microbiota were constructed using the FastTree method of QIIME 1.7.0 and the generated file was exported to the Interactive Tree of Life (iTOL) software [243] to view and format the constructed tree.

2.8. Statistical analysis.

Several statistical approaches were employed to identify taxa significantly associated with disease status and severity. Unless noted elsewhere, a Kruskal-Wallis test with post hoc Dunn's test was performed to compare the relative abundance of taxa as a function of disease status (CD vs. UC vs. control) and disease severity (mild vs. moderate vs severe). A Bonferroni correction was employed to account for multiple hypotheses. The relative abundance of the taxa identified were also analyzed by principal coordinate analysis (PCoA) and partial least squares discriminant analysis (PLS-DA). The PLS-DA models were validated by permutation tests. All statistical analyses were performed using XLSTAT (Addinsoft, NY) and/or GraphPad Prism version 6 (GraphPad, La Jolla, CA) softwares.

2.9. Prediction of the gut microbial functionalities in control, CD and UC groups.

High quality reads obtained after quality filtering and barcode sorting of Illumina generated reads were clustered into operational taxonomic units (OTUs) through QIIME 1.8.0 [236] via a closed-reference OTU picking workflow against the Greengenes reference set (release May, 2013) based on average percentage of identity of 97%. The resulting OTU

table was used for prediction of microbial functional profiles using the PICRUST algorithm [244] at the level of KEGG modules with known functions. The metabolic pathways identified were statistically compared among CD, UC and control groups using Kruskal-Wallis test followed by Dunn's post hoc test.

2.10. Quantitative polymerase chain reaction (qPCR)-validation of sequencing results.

The quantification of *A. parvulum, Fusobacterium nucleatum, Veillonella* genus and sulfate reducing bacteria relative to the total bacteria was determined by conducting relative quantitative PCR on the extracted metagenomic DNA using the Applied Biosystems 7300 DNA analyzer and primers listed in table 2.1. Each sample was tested in duplicate in a total volume of 25 µl per reaction. Fifty ng of template DNA was added to a reaction mixture containing 0.3 µM of each primer (0.5 µM for 16S rRNA universal primers) and 1x QuantiFast SYBR Green PCR master mix (Qiagen, Hilden, Germany). The amplification conditions were 5 min at 95°C followed by 40 cycles of 95°C for 10 sec and 66°C for 1 min with data collection at the second step of each cycle. Ct values were then extracted using the Applied Biosystems 7300 sequence detection software versions 1.3.1 and the relative abundance of each taxon was calculated as a ΔCt value (taxon Ct– universal 16S rRNA Ct). To validate the specificity of Apar-711F and Aparv-881R, fresh PCR amplicon from the total DNA extracted from two different mucosal aspirates was cloned using the TOPO® TA cloning vector (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. Next, the plasmid containing the 16S rRNA gene fragment was extracted from 6 different clones by the QIAprep Spin Miniprep kit (Qiagen, Hilden, Germany) followed by capillary sequencing using M13F and M13R primers.

Table 2.1: Primers used in qPCR validation of sequencing results.

Gene	Primer ID	Sequence	Reference
16S rRNA	16SUNIV-F 16SUNIV-R	5'-GGTGAATACGTTCCCGG-3' 5'-TACGGCTACCTTGTTACGACTT-3'	[245]
16S rRNA	Aparv-711F Aparv-881R	5'- GGGGAGTATTTCTTCCGTGCCG -3' 5'- CTTCACCTAAATGTCAAGCCCTGG -3'	This Study
16S rRNA	Fnucl-F Fnucl-R	5'-GGATTTATTGGGCGTAAAGC-3' 5'-GGCATTCCTACAAATATCTACGAA-3'	[246]
16S rRNA	Veilgen-F Veilgen-R	5'-AYCAACCTGCCCTTCAGA-3' 5'-CGTCCCGATTAACAGAGCTT-3'	[247]
Dissimilatory Sulfite Reductase (DSR)	DSR1-F DSR1-R	5'-ACSCACTGGAAGCACGGCGG-3' 5'-GTGGMRCCGTGCAKRTTGG-3'	[248]

2.11. Dextran Sodium Sulfate (DSS)-induced acute colitis model.

Four to six weeks old female Swiss Webster germ free mice raised in a germ free isolator were inoculated via gastric gavage with $1x10^8$ CFUs of either *Atopobium parvulum ATCC 33793* or *Bifidobacterium longum* subsp. *longum ATCC 55813* (n=5 mice per group). Colitis was then induced and development of inflammation was assessed as previously described [249, 250]. Briefly, the mouse drinking water was replaced with 3.5% DSS (36,000 – 50,000 MW, MP Biomedical) in sterile double distilled water for 5 days. On day 6 of DSS treatment, the mice were euthanized by CO_2 asphyxiation and the intestine was removed. The colon length from cecum to rectum was measured and the colon weight was determined. The colon, cecum and ileum were then opened longitudinally and flushed with

PBS to remove the intestinal contents. Half cm long segments of each site were used to determine intestinal colonization by the inoculated bacterium using a serial dilution viable count approach on fastidious anaerobic agar (Lab M) and tryptic soy agar (BD Biosciences) for *A. parvulum* and *B. longum*, respectively. Another 0.5 cm^2 piece was utilized for assessment of cytokine levels as described below. The remaining section from each site was Swiss rolled and immediately fixed in 10% buffered formalin. Five µm sections were embedded in paraffin and stained with hematoxylin and eosin. The severity of the inflammation was scored from zero (for healthy sections) to 5 (for severe inflammation) in a blinded fashion by two different investigators.

2.12. Assessment of cytokines in the intestinal tissues of DSS-treated mice.

The release of cytokines and inflammatory markers by intestinal tissues was measured as previously described [251]. Two 25 mm^2 explants from both colon and cecum were gently aspirated with PBS followed by shaking in 1 ml RPMI-1640 (Sigma-Aldrich) supplemented with 50 µg/ml gentamicin (Acros Organics) at room temperature for 30 minutes. Afterwards, the intestinal sections were transferred to 500 µl RPMI-1640 supplemented with 5% heat inactivated Fetal Bovine Serum (HI FBS, Invitrogen, Carlsbad, CA), 50 µg/ml gentamicin, and 1% penicillin/streptomycin (Sigma-Aldrich) and incubated at 37°C for 24 hours in a 5% CO_2 containing atmosphere. The supernatant was then collected and stored at –20°C until the measuring of the concentration of 32 cytokines and inflammatory markers via Multiplexing LASER Bead Technology (Eve Technologies Corporation, Alberta, Canada). To ensure reproducibility, cytokines released by two different segments from each location were investigated and normalized by the surface area of each segment. Finally, a Mann-Whitney U test was applied to differentiate the level of each marker among the tested groups.

2.13. *Il-10*$^{-/-}$ Mice experiments and tissue processing.

This experiment was conducted in collaboration with Dr. Christian Jobin and Dr. Marcus Mühlbauer at University of North Carolina, Chapel Hill. Germ-free SvEv129/C57BL6 *Il-10*$^{-/-}$;*NF-κB*^{EGFP} mice (8-12 weeks old, n=12) were transferred to specific pathogen free (SPF) conditions and mice from one cohort (n=6) were gavaged once weekly with *A. parvulum* (1x10^8 CFUs) for 6 weeks. *Atopobium parvulum* ATCC 33793 was grown in fastidious anaerobic broth (FAB) (Lab M, Canada). To investigate involvement of complex biota and H_2S in the development of colitis, we performed two subsequent experiments using 129/SvEv *Il-10*$^{-/-}$ mice. In the first experimental setting, gnotobiotic *Il-10*$^{-/-}$ mice (n=37) were randomized into 4 groups; 1: GF only (n=6), 2: GF + bismuth (III) subsalicylate (n=10); 3: *A. parvulum* (1x10^8 CFUs) (n=10) and 4: *A. parvulum* + bismuth (III) subsalicylate (n=11). Mice were euthanized after 6 weeks of mono-association. Bismuth (III) subsalicylate (Sigma-Aldrich, Saint Louis, MO) was incorporated into food (Teklan Global 18% Protein Rodent Diet) at a concentration of 7 g/kg (Harlan Laboratories, Madison, WI) and then irradiated for gnotobiotic experiments. Mice were fed with this diet starting one week before the colonization with *A. parvulum*. In the second experimental setting, gnotobiotic *Il-10*$^{-/-}$ mice (n=31) were transferred to SPF conditions and randomized into 4 groups; 1: SPF only (n=7), 2: SPF + bismuth (III) subsalicylate (n=8); 3: SPF plus *A. parvulum* (1x10^8 CFUs) (n=8) and 4: *A. parvulum* + bismuth (III) subsalicylate (n=8). Mice were euthanized after 6 weeks of weekly infection with *A. parvulum*. Bismuth (III) subsalicylate (Sigma-Aldrich, Saint Louis, MO) was incorporated into food (Teklan Global 18% Protein Rodent Diet) at a concentration of 7 g/kg (Harlan Laboratories, Madison, WI). Mice were fed with this diet starting at one week before the colonization with *A. parvulum*.

Tissue samples from the colon were collected for RNA and histology as previously described [252]. Histological images were acquired using a DP71 camera and DP Controller 3.1.1.276 (Olympus), and intestinal inflammation was scored as previously described [253]. The tissue was divided into 4 quarters, a score was given to each quarter separately and then added to generate a final inflammation score on a scale of 0 – 16.

2.14. Mouse endoscopy.

Colonoscopy was performed using a "Coloview System" (Karl Storz Veterinary Endoscopy) as described previously [254]. Mice were anesthetized using 1.5% to 2% isoflurane and ~4 cm of the colon from the anal verge from the splenic flexure was visualized. The procedures were digitally recorded on an AIDA Compaq PC.

2.15. Real time RT-PCR on mouse intestinal samples.

Total RNA from intestinal tissues was extracted using TRIzol (Invitrogen, Carlsbad, CA) following the manufacturers protocol. cDNA was reverse-transcribed using M-MLV (Invitrogen, Carlsbad, CA) and mRNA expression levels were measured using SYBR Green PCR Master mix (Applied Biosystems) on an ABI 7900HT Fast Real-Time PCR System and normalized to *β-actin.* The primers used were as follows: *β-actin* (5'-TGGAATCCTGTGGCATCCATGAAAC-3' and 5'-TAAAACGCAGCTCAGTAACAGTCCG-3'), *cxcl1* (5'-GCTGGGATTCACCTCAAGAA-3' and 5'-TCTCCGTTACTTGGGGACAC-3'), *tnf* (5'- ATGAGCACAGAAAGCATGATC-3' and 5'-TACAGGCTTGTCACTCGAATT-3'), *il12p40* (5'-GGAAGCACGGCAGCAGCAGAATA-3' and 5'- AACTTGAGGGAGAAGTAGGAATGG-3'), *il-1β*_(5'-GCCCATCCTCTGTGACTCAT-3' and 5'-AGGCCACAGGTATTTTGTCG-3'), *il-17a* (5'-TCCAGAAGGCCCTCAGACTA-3' and 5'-ACACCCACCAGCATCTTCTC-3'). The PCR reactions were performed for 40 cycles

according to the manufacturer's recommendation, and RNA fold changes were calculated using the ΔΔCt method [167].

2.16. Statistical analyses of *Il-10*$^{-/-}$ mice results.

Unless specifically noted, statistical analyses were performed using GraphPad Prism version 6 (GraphPad, La Jolla, CA). Comparisons of mice studies were made with a nonparametric analysis of variance, and then a Mann-Whitney U test. Experiments were considered statistically significant if $p < 0.05$.

2.17. Proteomic analysis of biopsies using super-SILAC-based quantitative mass spectrometry.

Proteomic analysis of right colon (RC) biopsies was conducted by the members of Dr. Daniel Figeys' laboratory. Briefly, Human hepatic HuH7 cells (HuH-7), human embryonic kidney 293 cells (HEK-293) and human colorectal cancer 116 cells (HCT-116) were individually grown and labeled following stable isotope labeling by amino acids in cell culture (SILAC) approach and by using heavy SILAC label (Heavy form of amino acids). Complete incorporation of heavy SILAC-label into HuH7, HEK-293 and HCT-116 cells was confirmed prior to collection and lysis of each cell type in RIPA buffer. Biopsies were lysed in 4% SDS, 50 mM Tris-HCl (pH 8.0) supplemented with protease inhibitor cocktail (Roche, Mannheim, Germany) and homogenized with a pellet pestle prior to sonication. The lysates of biopsies as well as the 3 reference cell types were centrifuged at 10000 x g for 10 min. The protein-rich supernatant was isolated and protein concentrations were determined using the Bio-Rad DC Protein Assay. The proteins were processed using the Filter Aided Sample Preparation Method (FASP) as previously described [255] with some modifications. Briefly, 45 µg of protein from RC tissue lysate was combined with an equal amount of pooled proteins from heavy SILAC labeled reference cells (15 µg protein from each cell protein

lysate). The resulting protein mixtures were reduced with 20 mM dithiothreitol, followed by alkylation with 20 mM iodoacetamide in 8 M urea, 50 mM Tris-HCL (pH 8.0). Excess SDS was removed by washing with 50 mM Tris-HCl pH 8.0 prior to tryptic digestion (TPCK Treated, Worthington) at 37°C overnight. Peptides were eluted by centrifugation and washing with 50mM Tris-HCl (pH 8.0), and then fractionated, using an in-house constructed SCX column with five pH fractions (pH 4.0, 6.0, 8.0, 10.0, 12.0). The fractionated samples were desalted using in-house C18 desalting cartridges and dried in a speed-vac. The resulting peptide mixtures were analyzed by high-performance liquid chromatography/ electrospray ionization tandem mass spectrometry (HPLC-ESI MS/MS) and the data were recorded with Xcalibur software (ThermoFisher Scientific, San Jose, CA). The generated raw data were processed and analysed using MaxQuant package (Version 1.2.2.5) against the decoy Uniport-human database. Statistical analyses were performed using Perseus (version 1.3.0.4). Comparisons of protein profiles were made with ANOVA with p values < 0.05 considered significant. Finally, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis was achieved using the DAVID Bioinformatics Resources (http://david.abcc.ncifcrf.gov/).

2.18. Total RNA extraction from mucosal aspirates for real-time qRT-PCR.

RNA integrity was preserved by adding an equal volume of RNAlater (Ambion) to the mucosal aspirates before freezing at -80°C. The frozen aliquot (2ml) was thawed on ice and the total RNA was extracted following the hot phenol protocol described previously [256-259]. Briefly, 4 ml of each sample in RNAlater were pelleted by centrifugation at 13,000 x g for 5 min at 4°C. The pellets were washed twice by resuspension in 50% RNAlater/PBS buffer and centrifugation at 13,000 x g for 5 min at 4°C. Next, the pellets were resuspended and lysed in 2 ml denaturing buffer (4M guanidium thiocyanate, 25mM

sodium citrate, 0.5% N-laurylsarcosine, and 1% N-acetyl cysteine, 0.1M 2-mercaptoethanol). The lysate was divided into 500 µl aliquots, to which 4 µl of 1M sodium acetate (pH 5.2) was added. Each aliquot was then incubated with 500 µl of buffer saturated phenol (pH 4.3) at 64°C for 10 minutes, with intermittent mixing. The tube was placed on ice for 10 min followed by centrifugation at maximum speed for 30 min at 4°C. One ml of chloroform was added to the aqueous layer and incubated for 15 minutes on ice, followed by centrifugation at 18,000 x g for 30 min at 4°C. Afterwards, RNA was precipitated from the aqueous layer by the addition of 1/10 volume 3M sodium acetate, 500 mM DEPC treated EDTA and 2 volumes of cold ethanol followed by overnight incubation at -80°C. Later, the RNA was pelleted by centrifugation at maximum speed for 30 min at 4°C, aspirated with 80% cold ethanol and resuspended in 100 µl of RNAse free dH_2O. The extracted RNA was treated twice with Dnase I (Epicentre) followed by PCR amplification using the 16S rRNA universal primers; Bact-8F and 1391-R [101], to confirm the absence of genomic DNA. The quality and the quantity of the extracted RNA were determined by NanoDrop 2000 spectrophotometer (Thermoscientific) and confirmed by BioRad's Experion StdSens RNA system according to the manufacturer's description and stored at -80°C until use.

2.19. qRT-PCR quantification of hydrogen sulfide detoxification genes expression level.

The quantification of the expression level of *TST* (Thiosulfate Sulfurtransferase), *SQRDL* (Sulfide Quinone Reductase Like) and *COX4-1* (Cytochrome C oxidase subunit IV isoform 1) relative to *hGAPDH* (Glyceraldehyde-3-Phosphate Dehydrogenase) genes was determined using the Applied Biosystems 7300 DNA analyzer and Quantitect SYBR Green RT-PCR kit (Qiagen, Hilden, Germany). The primers used were either designed using the NCBI Primer-BLAST tool [260] or extracted from the literature as mentioned in table 2.2.

The specificity of the primers was confirmed by TOPO® TA cloning and capillary sequencing as previously mentioned in section 2.10. Each reaction contained 100 ng RNA template, 0.5 µM of each primer, 1x Quantitect SYBR Green RT-PCR master mix and 0.25 µl Quantitect RT-mix. The one step qRT-PCR conditions were 50°C for 30 min, 95°C for 15 min followed by 40 cycles of 15 sec at 94°C, 30 sec at 60°C and 30 sec at 72°C with data collection at the third step of each cycle. The amplification specificity was checked by the melting profile of the amplicon and 2% agarose gel electrophoresis. Ct values were then extracted using the Applied Biosystems 7300 sequence detection software versions 1.3.1. Ct values of *TST, SQRDL*, or *COX4* were normalized to the Ct values of *hGAPDH* generating ΔCt. Next, ΔΔCt was calculated by subtracting the average ΔCt of the control group from the ΔCt of each sample. The relative quantification was then calculated as 2^-ΔΔCt as mentioned previously [167].

Table 2.2: Primers used in qRT-PCR.

Gene	Primer name	Sequence	Reference
hGAPDH	hGAPDHF	5'- AGAAGGCTGGGGCTCATTTG -3'	[261]
	hGAPDHR	5'- AGGGGCCATCCACAGTCTTC -3'	
TST	hTSTF	5'- TCTCAAGGGCGGTTCCTGGG -3'	This study
	hTSTR	5'- CGTACACGGCCACATCAGGC -3'	
SQRDL	SQRDL951F	5'- AATGCCATCTTCACCTTCCC -3'	This study
	SQRDL1105R	5'-TTAACCCCGAAAATGGCTCC-3'	
COX4-1	hCOX497F	5'- GTGGGCGGTGCCATGTTCTT -3'	This study
	hCOX757R	5'- GCATGGAGTTGCATGGCGGT -3'	

CHAPTER 3.Results

3.1. Demographic and clinical characteristics of the pediatric cohort employed.

This study involved the enrollment, detailed assessment, and biological sampling of 217 pediatric subjects (105 CD, 38 UC, and 74 controls; Table 3.1). A positive family history of one or more first degree relatives (parents or siblings) with CD, UC or IBDU was confirmed in 28 subjects (13 controls, 11 CD and 4 UC; Table 3.1). The diagnosis of these subjects was confirmed by histological examination, and the severity of inflammation was expressed using the Pediatric Crohn's Disease Activity Index/Pediatric Ulcerative Colitis Activity Index (PCDAI/PUCAI; Appendix II). Most IBD subjects had active inflammatory luminal disease involving the terminal ileum and/or the colon +/- perianal disease. As a function of PCDAI, CD subjects were categorized into inactive, mild, moderate and severe subgroups (n= 6, 25, 26 and 48, respectively; Table 3.1). The disease severity of UC subjects also ranged from inactive, mild, moderate to severe inflammation (n=3, 4, 17 and 14, respectively; Table 3.1). The location of CD inflammation was either ileal (n=32), colonic (n=21), ileocolonic (n=50) or jejunal/proximal ileal inflammation (n= 1; Table 3.1), while the extent of UC inflammation involved either ulcerative proctitis (only the rectum (E1); n=3), left sided UC (distal to the splenic flexure (E2); n=3), extensive UC (distal to the hepatic flexure (E3); n=7) or pancolitis (the whole colon (E4); n=25; Table 3.1). Finally, the age and gender characteristics of the cohort are presented in table 3.1.

Table 3.1: Metadata statistics for the subjects involved in the current study.

Description*		**Control**	**CD**	**UC**
Subjects		74	105	38
Positive Family History	CD	7	9	0
	UC	5	1	4
	IBD-U	1	1	0
Age				
	Range (year)	3-18	3-17	4-18
	Median Age, years (IQ Range)	15 (11-16)	14 (11.5-15)	15 (13-17)
Gender				
	Female	39	36	18
	Male	35	69	20
Disease Activity (PCDAI/PUCAI)**				
	Inactive	-	6	3
	Mild	-	25	4
	Moderate	-	26	17
	Severe	-	48	14
Disease Location***				
		-	Upper 1	E1 3
		-	Ileal 32	E2 3
		-	Colonic 21	E3 7
		-	Ileal-Colonic 50	E4 25

*CD; Crohn's disease; UC; ulcerative colitis. **PCDAI/PUCAI; Pediatric CD/UC Activity Index. ***E1; ulcerative proctitis; E2; left sided UC; E3; Extensive UC and E4; pancolitis.

3.2. Differences in microbial signature identified from mucosal aspirates, mucosal biopsies and stools.

3.2.1. Diversity of the gut microbiota from different sampling sources.

The microbiota retrieved from stool samples are known to exhibit higher diversity as compared to biopsy collected microbiota [91]. In order to compare the microbial diversity in mucosal aspirates to that of stool and mucosal biopsies, we collected these 3 sample types from 4 children with CD. The mucosal biopsies were collected from both inflamed and non-inflamed tissues. The microbial diversity in these samples was determined by sequencing the hypervariable V6 region of the 16S rRNA gene. A total of 9,704,764 high quality reads obtained after quality filtering and were assigned successfully to 2136 OTUs (97% similarity) using the closed reference OTU picking approach of QIIME 1.7 [236]. Good's coverage score was above 99% for each sample. The average numbers of observed OTUs per category were 616.6, 621.8, 640.75 and 748.8 for colonic aspirates (COasp), colonic biopsies from non-inflamed (BCON) and inflamed tissues (BCOA) and stool samples, respectively. Chao1 and Shannon indices were calculated using 299,550 randomly selected reads per sample to estimate microbial richness and diversity, respectively. We statistically compared the calculated Chao1 and Shannon indices using Freidman paired test followed by Dunn's post hoc test. No significant difference was found in the richness of the gut microbiota among different sample types. However, the stool microbiota showed a significantly higher diversity as compared to the microbiota from non-inflamed biopsies ($p < 0.05$; Figure 3.2.1 A-B). To gain a general overview of the gut microbiota similarity between different sample types, we conducted a principle coordinate analysis (PCoA) based on the un-weighted UniFrac distances. The PcoA separated the samples into 2 clusters with the two component axes accounting for 24.08% and 16.56% of the total variation; One cluster contained the

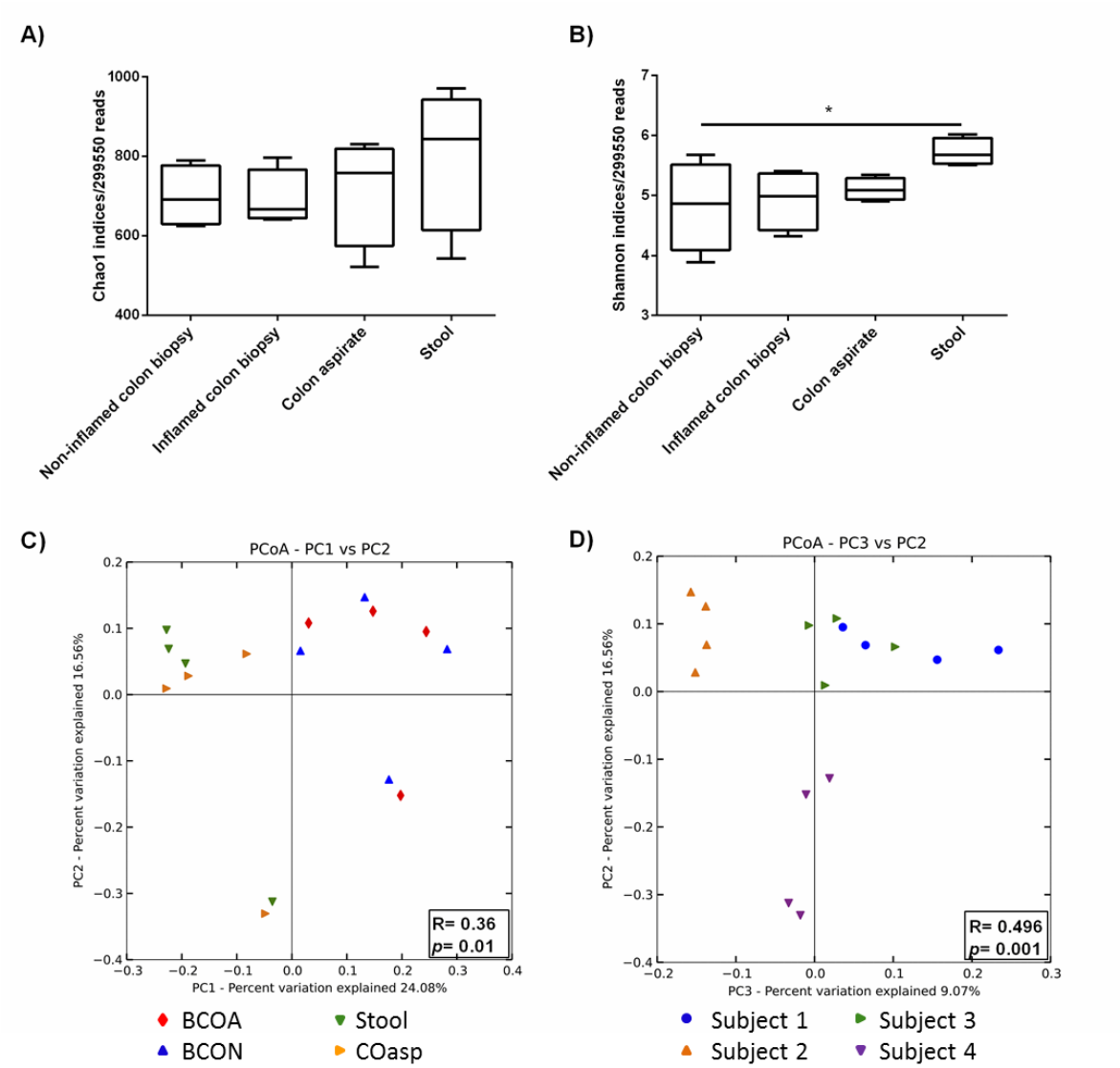


Figure 3.2.1: Diversity of the gut microbiota identified from mucosal aspirates (COasp), mucosal biopsies (BCOA for inflamed tissues and BCON for non-inflamed tissues) and stool samples. A, Richness of the gut microbiota as expressed by Chao1 indices of each sample from different sampling methods (n=4 per group). **B,** Diversity of the gut microbiota as expressed by Shannon indices of the same samples as in panel A. 299,550 reads were randomly extracted from the total sequences of each sample by QIIME 1.7 and used to calculate the diversity metrics. The horizontal black line represents the mean. Freidman test followed by Dunn's multiple comparison were used for paired statistical comparisons (*p< 0.05). **C-D,** Principle Coordinate Analysis (PCoA) based on the un-weighted UniFrac distances between the different samples was conducted using QIIME 1.7. Analysis of similarity (ANOSIM) was used to test for dissimilarity between groups using 999 permutations, where R = 1 corresponds to a maximum dissimilarity. C, PCoA of the identified OTUs colored by the sample type. D, PCoA of the identified OTUs colored by the source subject.

microbiota from both the inflamed and non-inflamed biopsies, and the other the mucosal aspirates and stool samples (Figure 3.2.1 C). Additionally, the microbiota clustered by the source subject with the two component axes accounting for 9.07% and 16.56% of the total variation (Figure 3.2.1 D). These results indicate that the gut microbiota identified from mucosal biopsies are substantially different from the microbiota sampled by stools and mucosal aspirates. Furthermore, no substantial difference in the diversity was evident between the gut microbiota from inflamed and non-inflamed tissues. Finally, the subject-based variation within stool samples or colonic aspirates is greater than the variation between the two approaches.

In order to reduce the impact of interpersonal variabilities on the identified microbiota, we identified the core microbiota that are conserved amongst 75% of the individuals within each sample type. The core microbiota indicates the diversity and structure of the microbiota conserved between samples within a certain group. A total of 613 OTUs were identified as a core microbiota from different sampling methods, which represent 95.4±2.7% of the total microbial population. The smallest core microbiota was identified from BCON followed by COasp, BCOA and stool samples with 321, 355, 407 and 440 core OTUs, respectively (Figure 3.2.2 A-B). Two hundred and seven core microbial OTUs were shared by the 4 sample types (Figure 3.2.2 A). Stool samples harbored the highest number of unique core OTUs (94 OTUs) followed by BCOA, COasp and BCON (61, 24 and 18 OTUs, respectively; Figure 3.2.2 A). However, no significant difference in either the richness or diversity of the identified core microbiota was detected between the different sample types (Figure 3.2.2 C-D).

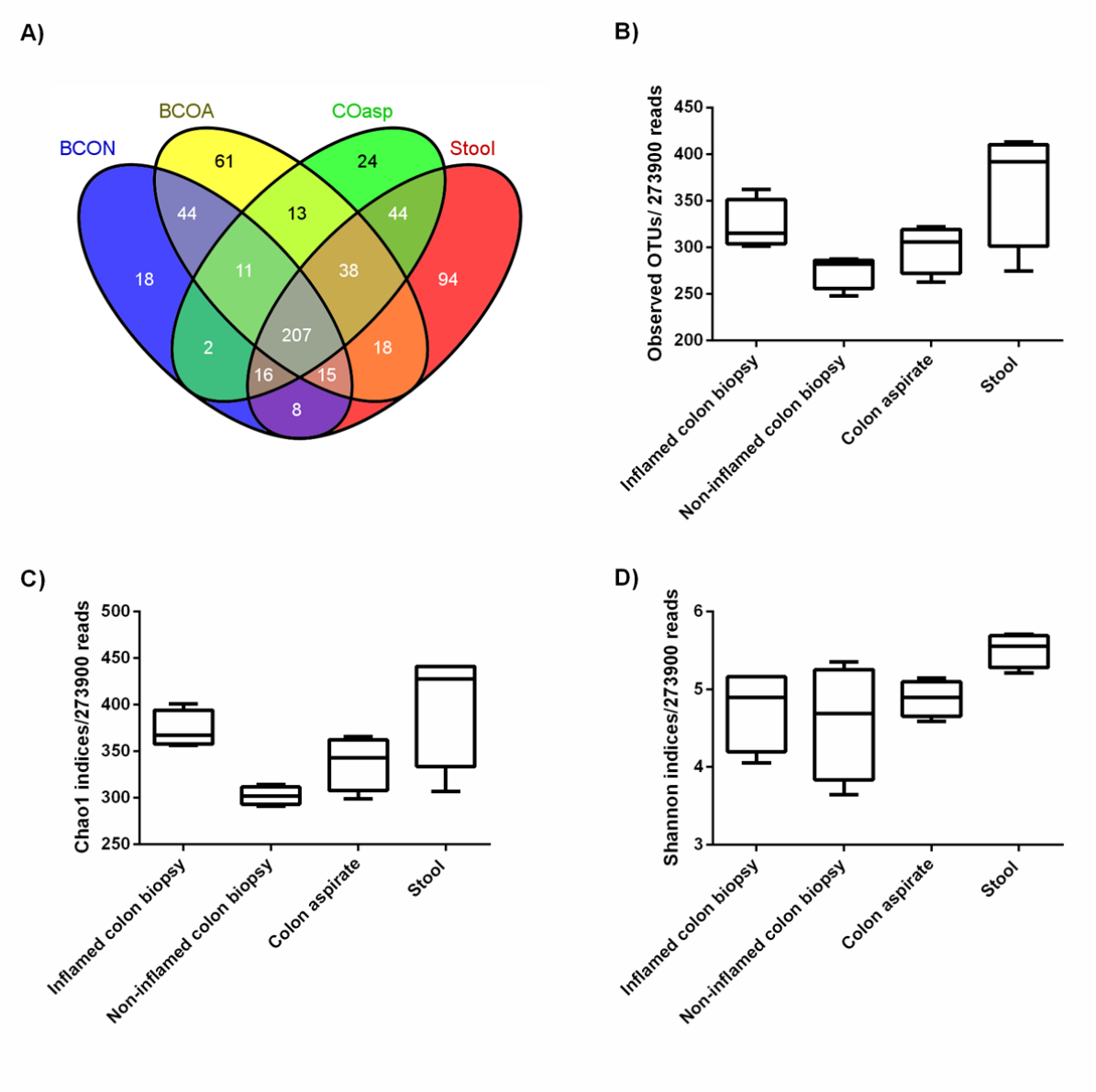


Figure 3.2.2: Diversity of the core gut microbiota (0.75) identified from mucosal aspirates (COasp), mucosal biopsies from inflamed (BCOA) and non-inflamed tissues (BCON) and stool samples. A, Venn diagram showing the number of core OTUs (0.75) that is either unique for each sample type or shared between two or more types. **B,** The number of the core OTUs identified in each sample of differing sampling sources (n=4 per group). **C,** Richness of the core microbiota as expressed by Chao1 indices of the same samples as in panel B. **D,** Diversity of the core microbiota as expressed by Shannon indices of the same samples as in panel B. Randomly selected 273,900 reads from the sequences belonging to the core OTUs of each category were used to calculate the diversity metrics. The horizontal black line represents the mean. Statistical comparison was conducted using Freidman test followed by Dunn's multiple comparison.

3.2.2. Taxonomic composition of the gut microbiota from biopsies, mucosal aspirates and stool samples.

Prior literature illustrated that the microbiota harbored by stool samples differs from the resident microbiota at the intestinal mucosa [139]. In order to investigate the similarity of the microbial structure between mucosal aspirates and either stool samples or mucosal biopsies, we statistically compared the microbial taxa identified from each sample type. A total of 15 phyla were identified from all of the samples. While fourteen phyla were detected in BCON, we identified 12 phyla in BCOA, 11 phyla in COasp and only 10 phyla in stool samples. Firmicutes and Bacteroidetes, as expected, were the two predominant phyla representing 74-93% of the total microbiota (Figure 3.2.3 A-B). The average relative abundance of Firmicutes±SEM were 31.9±5.7%, 35.68±6.2%, 52.23±2.6% and 53.2±3.4%, while the relative abundance of Bacteroidetes were 42.5±7.4%, 41.5±7.8%, 37.8±4.9% and 40.6±2.1% in BCOA, BCON, COasp and stool samples, respectively (Figure 3.2.3 A-B). Other phyla included Acidobacteria, Actinobacteria, Chloroflexi, Cyanobacteria, Fusobacteria, Nitrospirae, Planctomycetes, Proteobacteria, Spirochetes, TM7, Tenericutes, Thermi and Verrucomicrobia. A Freidman test followed by a Dunn's post hoc test was used to compare the relative abundance of different taxa. Firmicutes, Proteobacteria, Acidobacteria, Cyanobacteria and Planctomycetes were the only phyla that showed significant differences in their abundance among the microbiota of BCOA, BCON, COasp and/or stool samples (Freidman p <0.05; Figure 3.2.3 C). Firmicutes were significantly depleted along with the enrichment of Proteobacteria in the microbiota of BCOA or BCON compared to the stool samples. While Acidobacteria was significantly higher in the BCOA microbiota in comparison with stool samples, Cyanobacteria was significantly enriched in

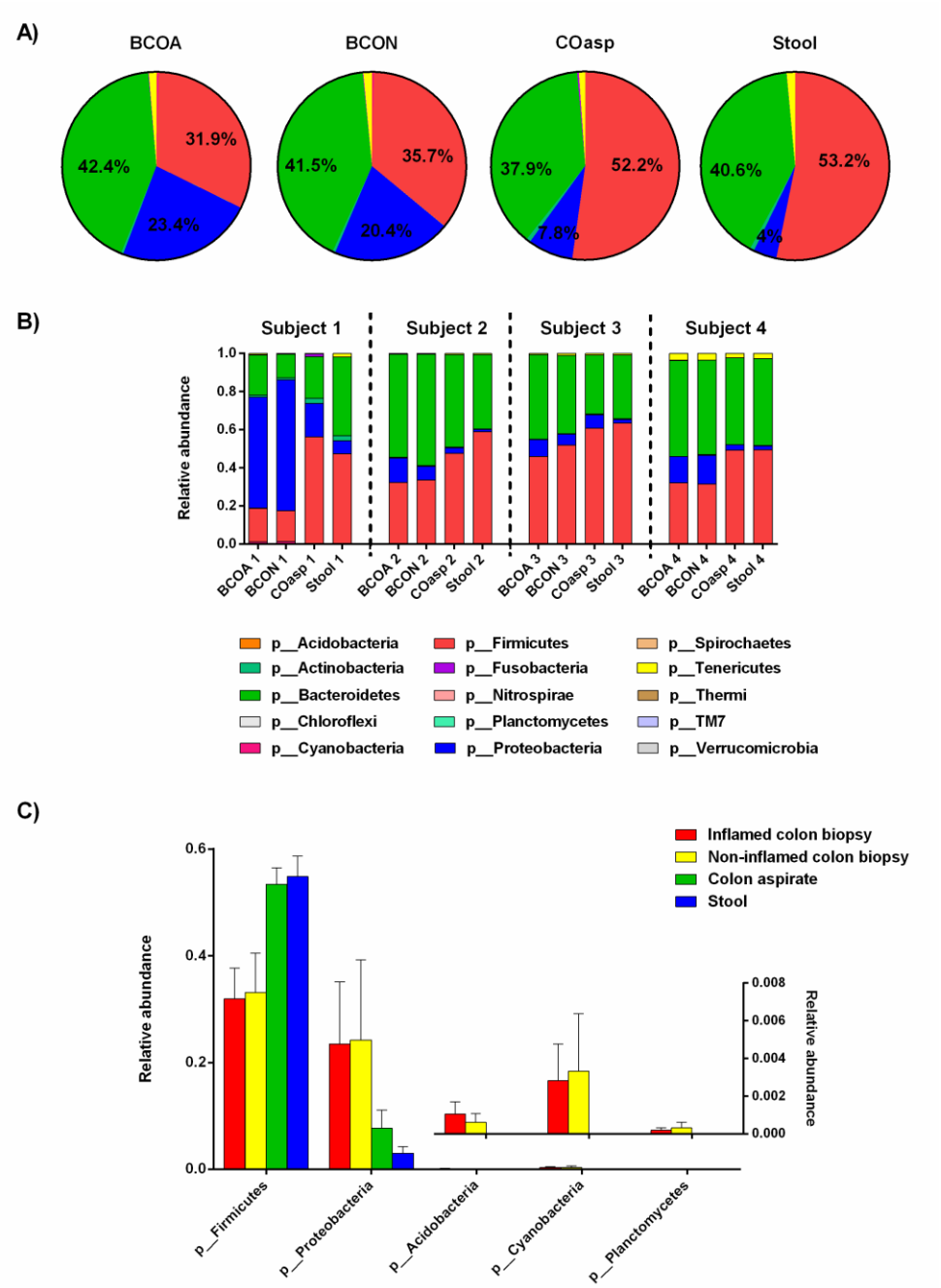


Figure 3.2.3: Relative abundance of bacterial phyla from colon biopsies, colon aspirates and stool samples. A, Average relative abundance of dominant phyla from colon biopsies of inflamed (BCOA; n=4) and non-inflamed tissues (BCON; n=4), colon aspirates (COasp; n=4) and stool samples (n=4). **B,** Relative abundance of major phyla in each sample from the individual subjects. **C,** Relative abundance of differentially abundant phyla from panel A and B. The bar represents the mean±SEM of each group. A Freidman test followed by Dunn's post hoc test was used for statistical analysis. The inset histogram shows the relative abundance of taxa present at low abundance on the right y axis.

the BCON microbiota relative to the COasp (Figure 3.2.3 C). At lower taxonomic levels, Propionibacteriaceae (0.19±0.07% vs $7.7x10^{-5}$±$5.12x10^{-5}$%), Gemellaceae (0.05±0.03% vs $7.8x10^{-3}$±$6.5x10^{-3}$%), Vibrionaceae (0.19±0.14% vs $2.3x10^{-2}$±$1.9x10^{-2}$%), *propionibacterium* (0.19±0.07% vs $7.7x10^{-5}$±$5.12x10^{-5}$%), *Gemella* (0.05±0.03% vs $7.8x10^{-3}$±$6.5x10^{-3}$%) and *Pseudomonas* (0.04±0.03% vs $4.6x10^{-4}$±$3.5x10^{-4}$%) were significantly enriched in the BCON microbiota as compared to the fecal microbiota, while the relative abundance of Ruminococcaceae (12.46±2.3% vs 18.1±4%) and *Blautia* (1.87±0.9% vs 4.1±1.3%) were significantly lower in the BCON microbiota than in the stool microbiota (Figures 3.2.4 and 3.2.5). In comparison with the COasp microbiota, the relative abundance of Staphylococcaceae (0.07±0.03% vs 0.0013±0.001%) and its genus *Staphylococcus* (0.07±0.03% vs 0.0013±0.001%) were increased while that of Lachnospiraceae (18.96±5.43% vs 34.07±3.47%) was decreased in the BCON microbiota (Figures 3.2.4 and 3.2.5). On the other hand, the BCOA microbiota showed increased relative abundance of *Solibacteres* (0.09±0.05% vs 0.00±0.00), Betaproteobacteria (18.58±7.2% vs 2.6±1.0%), Solibacteraceae (0.09±0.05% vs 0.00±0.00) and *Burkholderia* (0.04±0.03% vs 0.00±00%) as compared to the stool microbiota; Burkholderiaceae (10.59±8.77% vs 0.0031±0.0014%), *Cupriavidus* (9.6±5.3% vs 0.0024±0.0014%) and *Commamonas* (0.3±0.16% vs $3.2x10^{-5}$±$3.2x10^{-5}$%) in comparison with the COasp microbiota (Figures 3.2.4 and 3.2.5).

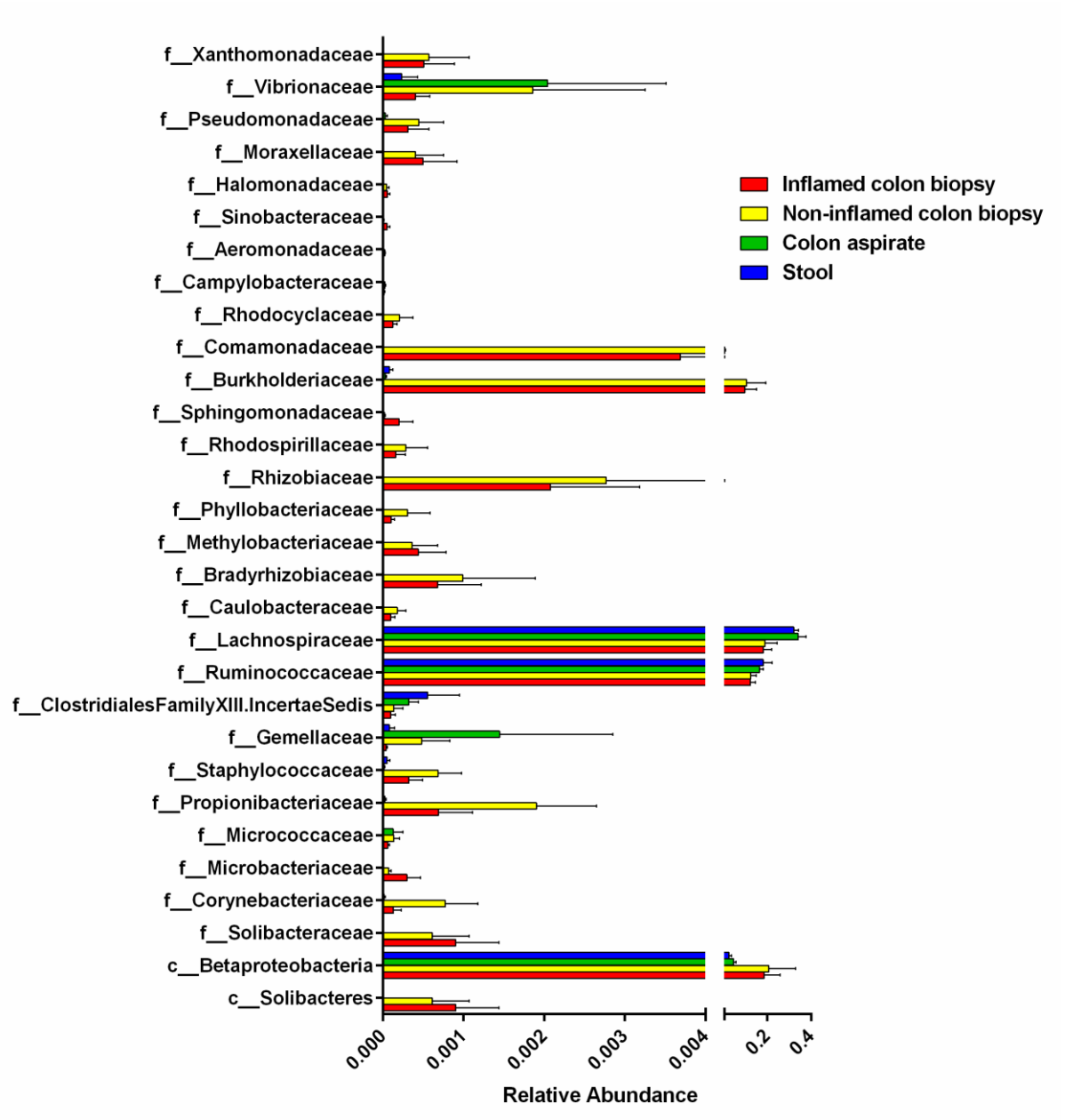


Figure 3.2.4: Bacterial classes and families that vary significantly in abundance in at least one of the three sample types (Biopsies, stool and aspirates). The figure represents the relative abundance of bacterial classes and families that are differentially abundant between the microbiota from different sampling sources (n= 4 per group). The relative abundance of all bacterial classes and families identified by QIIME 1.7 analysis were subjected to Freidman test followed by Dunn's post hoc test.

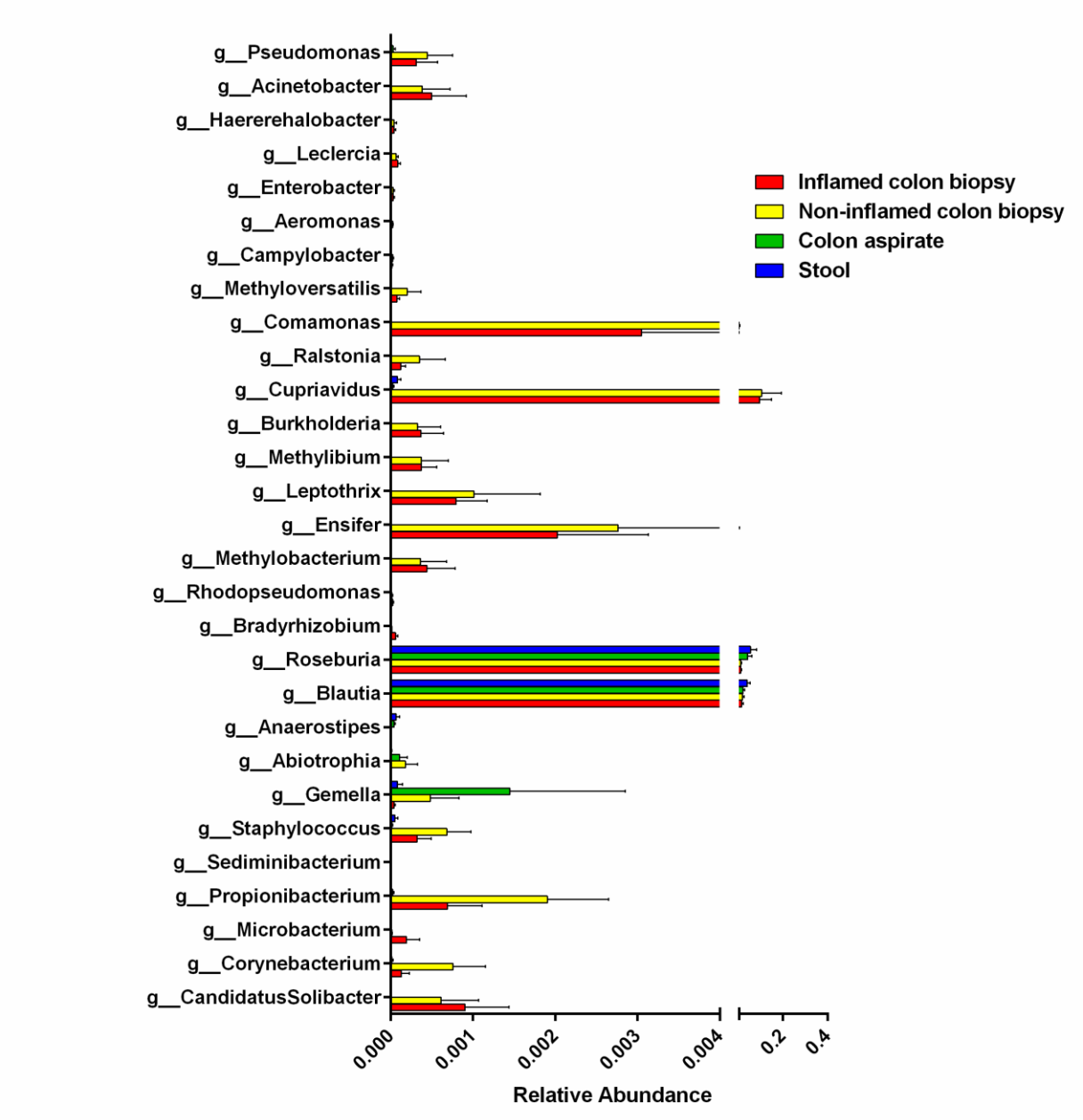


Figure 3.2.5: Bacterial genera that vary significantly in abundance in at least one of the three sample types (Biopsies stool and aspirates). The figure represents the relative abundance of bacterial genera that are differentially abundant between the microbiota from different sampling sources (n= 4 per group). The relative abundance of all bacterial genera identified by QIIME 1.7 analysis were subjected to Freidman test followed by Dunn's post hoc test.

3.3. Microbial diversity and structure along the gut of healthy children.

3.3.1. Gut microbial diversity of healthy children.

Several previous studies have characterized the gut microbiota of adolescent children [262-265]. However, these studies relied on stool samples to collect the gut microbiota. It is known that the microbiota harbored by stool samples differ from what is embedded within the intestinal mucosa in terms of both diversity and structure [91, 139]. Therefore, we aimed to use the mucosal aspirates to characterize the diversity and structure of the gut microbiota at the left colon (LC), right colon (RC) and terminal ileum (TI) of healthy children from the age range of 3 to 18 years old Using Illumina HiSeq approach. A total of 66,900,789 high quality reads were obtained by sequencing the V6 region of the 16S rRNA gene from 70 samples representing the three different intestinal locations (n=34, 21 and 15 for LC, RC and TI, respectively). Based on 97% sequence similarity, these reads were assigned to 5998 OTUs. No significant difference was identified in either the richness or diversity of the microbial communities along the intestinal tract (Figure 3.3.1 A and D).

To assess the microbial diversity as a function of age, we categorized the samples from each intestinal location into 3 age groups: Age group A for samples from children less than 10 years old; Age group B for children 11-14 years old and age group C for 15-18 years old children. No significant difference was detected in the richness or the diversity amongst the 3 age groups at any tested intestinal location (Figure 3.3.1 B and E). Furthermore, no difference in the microbial diversity was observed between male or female children (Figure 3.3.1 C and F). To compare the overall composition of the gut microbiota along the intestinal tract, we conducted a principle coordinate analysis (PCoA) based on the un-weighted UniFrac distances. In addition, analysis of similarity (ANOSIM) between groups was applied

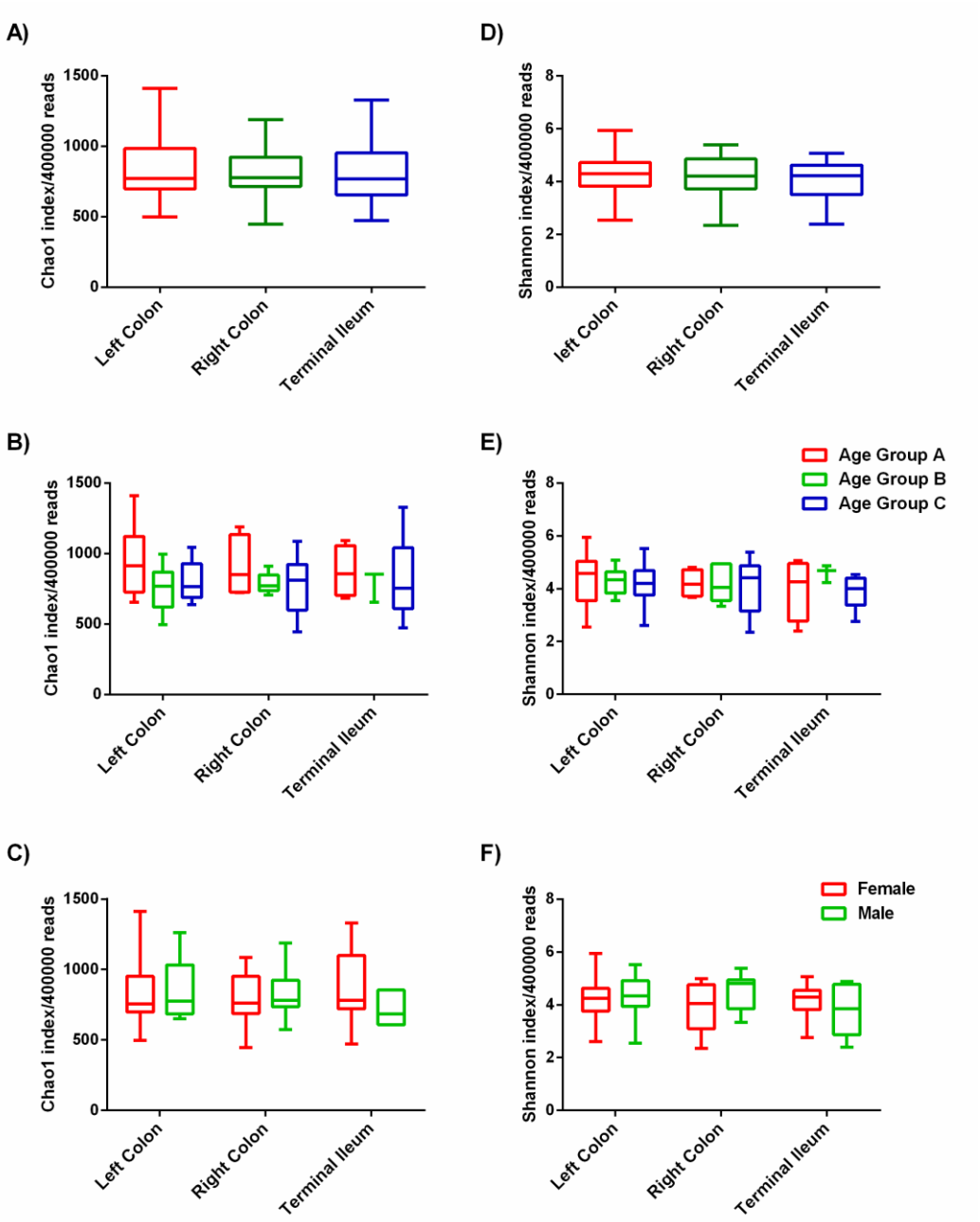


Figure 3.3.1: Gut microbiota exhibit similar diversity along the intestinal tract of non-IBD children and between different age and gender groups. The richness and diversity of the gut microbiota were expressed by Chao1 and Shannon indices which were calculated using QIIME 1.7. **A,** Chao1 indices at the three tested intestinal locations (A; n=34, 21 and 15 for LC, RC and TI, respectively). **B-C,** Categorization of the samples in panel A as a function of age and gender. **D-F,** Shannon indices for the same categories as in panels A-C. A Kruskal-Wallis test followed by Dunn's multiple comparisons were applied for statistical comparison. Age group A is for children < 10yrs old, age group B is 10-14yrs old and age group C ranges from 15 to 18yrs old.

using 999 permutations, where R = 1 is the maximum dissimilarity. The microbiota from different samples clustered based on the source subject on the first two components axes accounting for 12.07% and 8.52% of the total variation (Figure 3.3.2 A; ANOSIM R=0.905 and p= 0.001). On the other hand, the microbial communities did not separate based on the intestinal location, age group or gender (Figure 3.3.2 B-D; R= 0.008, 0.104 and 0.051, respectively). This result indicates that the interpersonal microbial variations are greater than the differences between the sampling locations, ages and/or the gender of healthy children.

3.3.2. Microbial structure along the length of the intestinal tract of non-IBD children.

In order to identify the differential taxa within each intestinal location, age group or gender, we statistically compared the relative abundance of each identified taxa between the tested groups. The microbiota of healthy children were characterized at the phyla level by a gradual decrease in the relative abundance of Bacteroidetes, as a result of depletion of Bacteroidaceae and Porphyromondaceae families, from LC to RC to TI with relative abundance±SEM of 36.3±2.9% at LC, 26.9±4.3% at RC and 24.5±4.9% at TI. In contrast, Firmicutes progressively increased, due to enrichment of Lachnospiraceae, from LC to RC to TI with relative abundance±SEM of 47.3±2.6% at LC, 56.4±4.4% at RC and 51.9±5.3% at TI (Figure 3.3.3). A Kruskal-Wallis comparison of the microbial relative abundance between LC, RC and TI identified 62 taxa that were differentially distributed along the intestine of non-IBD children (Appendix III). At other phylogenetic levels, *Brachybacterium*, Sanguibactraceae, *Pedobacter*, Streptococcaceae, Phylobacteriaceae, Rhizobiaceae, and Collowiaceae, Nocardiaceae, Bradyrhizobiaceae, *Chloroflexi*, Hyphomicroboaceae, Hallothiobacillaceae, *Aquimonas* and *Dyella* were enriched at the terminal ileum as compared

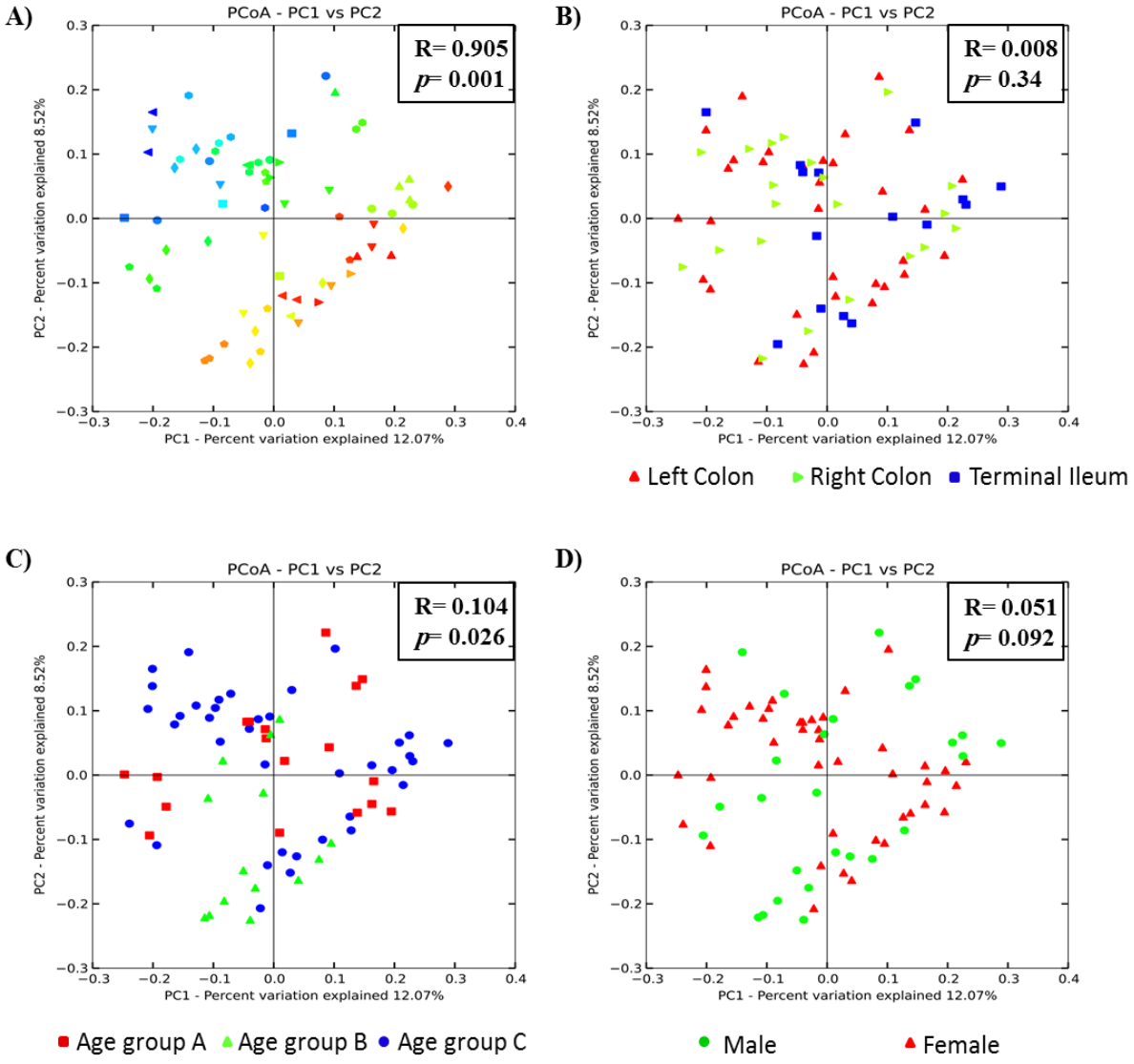


Figure 3.3.2: Phylogenetic beta-diversity analyses revealed individuality of microbial communities along the intestinal tract between different subjects. PCoA analyses based on un-weighted UniFrac distances were conducted using QIIME 1.7 to compare microbial communities at three intestinal locations (left colon; Right colon and Terminal ileum). The samples were colored either by source subject **(A)**, sampling site **(B)**, age category **(C)** or gender **(D)**. Analysis of similarity (ANOSIM) was used to test for dissimilarity between groups using 999 permutations with R = 1 indicating maximum dissimilarity. Age group A corresponds to children < 10yrs old, age group B to 10-14yrs old children and age group C to 15-18yrs old children.

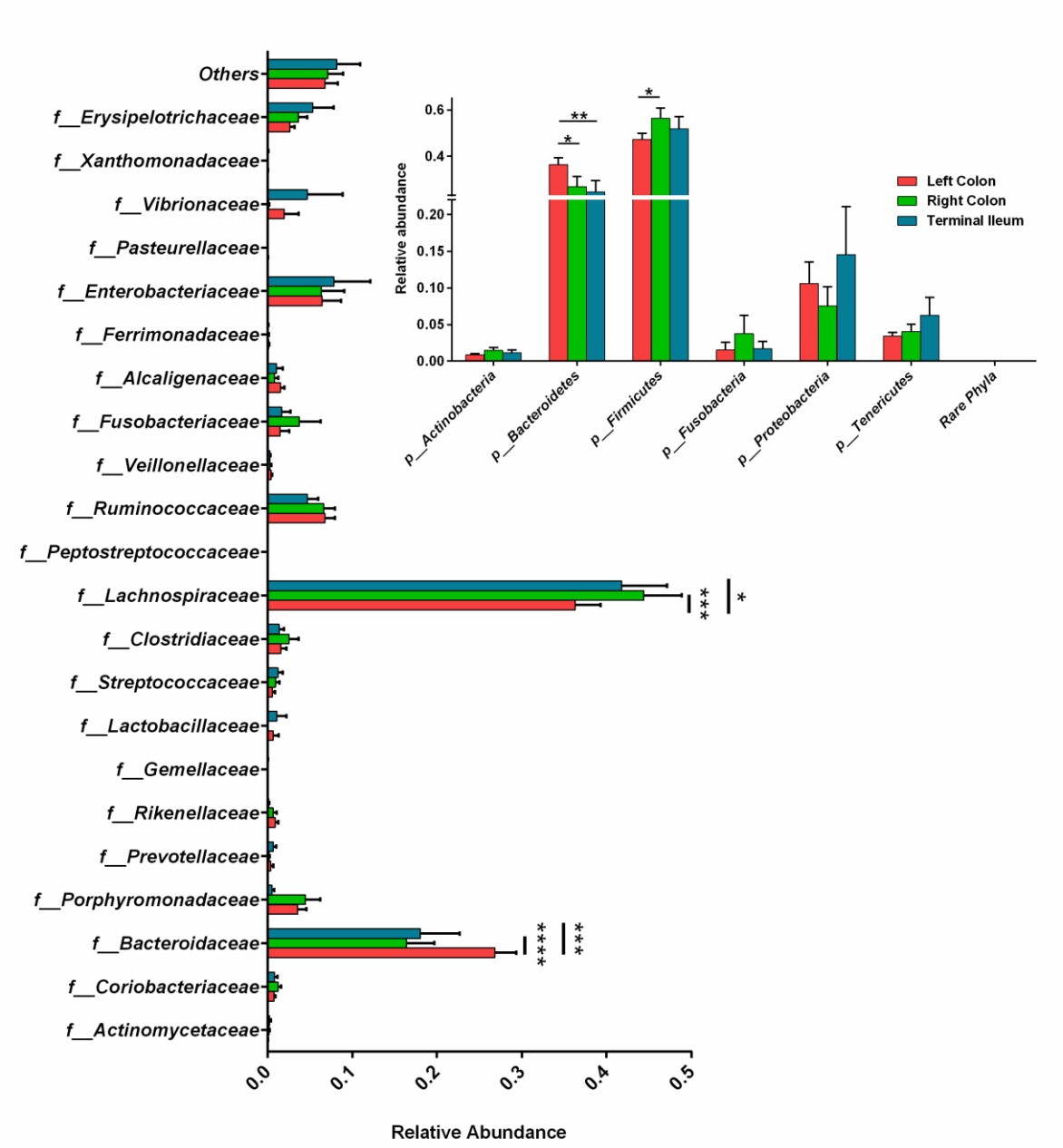


Figure 3.3.3: Microbial structure along the length of the intestinal tract of healthy children. The relative abundance of different phyla and families were generated using QIIME 1.7 and the means±SEM were plotted using GraphPad Prism 6. The figure shows the average relative abundance of the major phyla (top right panel) and families at the left colon, right colon and terminal ileum of healthy children. Two-way ANOVA followed by Bonferroni correction for multiple comparisons were used for statistical comparisons. *p<0.05, ***p<0.001 and ****p<0.0001.

to the RC and/or LC, *Bacteroides and Oceanobacillus* were more represented at the LC, and finally, Bacilli, Lactobacillales, Carnobacteriaceae and *Vagococcus* increased at the RC in comparison with the LC and TI microbiota (Appendix III).

To examine the development of the gut microbiota as a function of age, we divided the children into 3 age groups; group A for children under 10 years, group B for 11-14 years old and group C for children of 15-18 years old. The relative abundance of the taxa in these 3 groups were compared using a Kruskal-Wallis test. The major differences in the microbial structures were noticed in group B as compared to the other two groups at the three intestinal locations (Figure 3.3.4 A-C). At the LC, Flexibacteraceae, *Chloroplast*, *Streptophyta*, Lactobacillaceae, Peptococcaceae, *Proteus*, *Anaerococcus* and Rhodocyclaceae showed higher relative abundance in age group A as compared to groups B and/or C (Appendix IV). The RC microbiota was characterized by increased relative abundance of Aerococcaceae, *Subdoligranulum* and *Facklamia* in age group B as compared to group C, while the relative abundance of *Abiotrophia*, *Aerococcus*, *Enterococcus*, *Anaerococcus*, *Moryella*, *Oribacterium*, Alphaproteobacteria, *Enterobacter*, *Tatumella*, *Marinomonas* and *Thermomonas* decreased in age group B as compared to groups A and/or C (Appendix V). The terminal ileum microbiota showed the least changes across ages with *Odoribacter* and *Oscillospira* as the only taxa depleted in 11-14 years old children as compared to 15-18 years old children, and *Clostridium* and unclassified Rikenellaceae as the only taxa increased in children less than 10 years old as compared to 11-14 years old children (Appendix VI). Interestingly, we did not detect any differences in the microbial structure between male and female children.

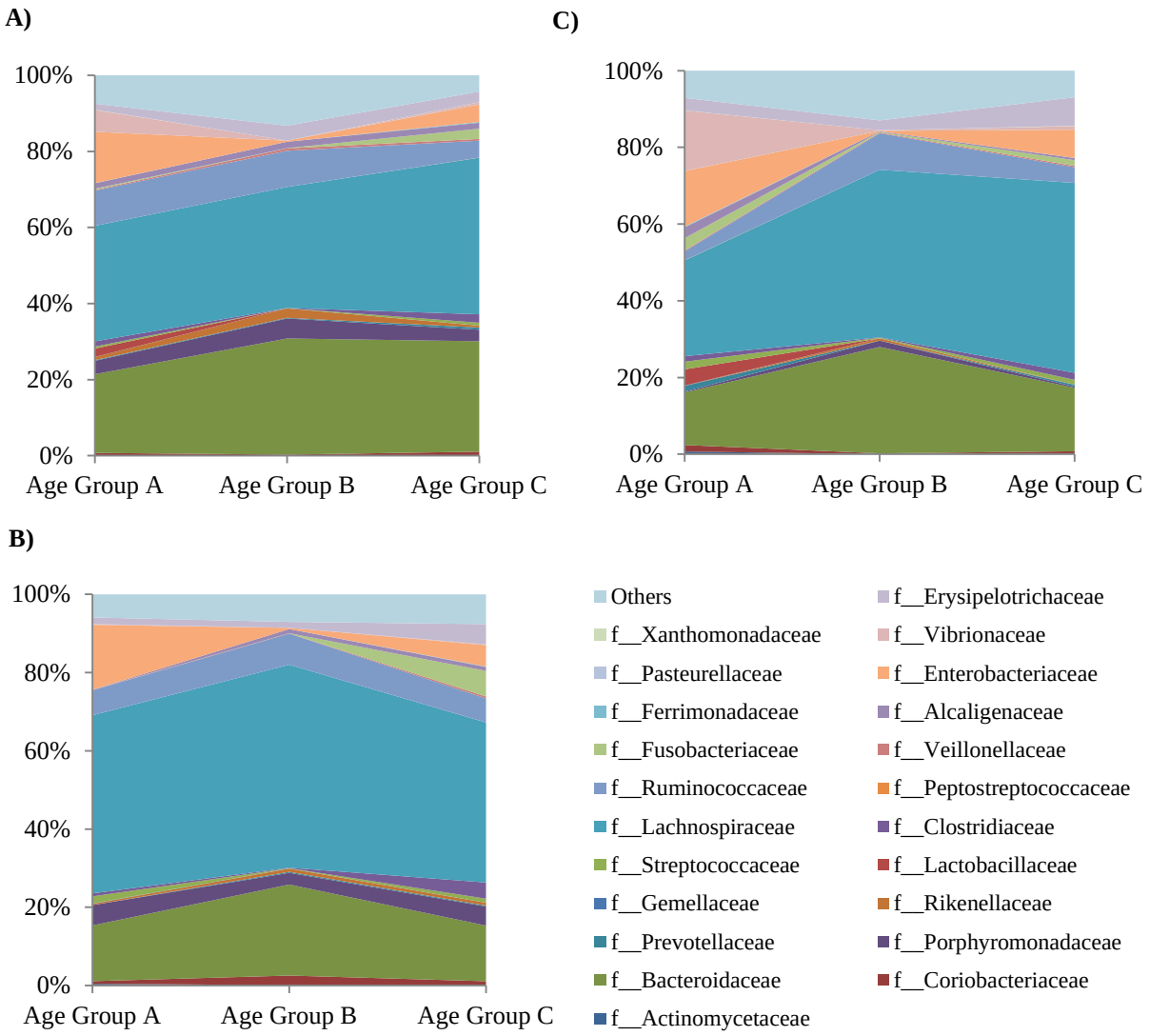


Figure 3.3.4: Microbial structure along the length of the intestinal tract of healthy children at different ages. The relative abundance of bacterial families were generated using QIIME 1.7. The figure shows the average relative abundance of predominant families across ages at the left colon **(A)**, right colon **(B)** and terminal ileum **(C)** of healthy children. Age group A corresponds to children < 10yrs old, age group B corresponds to 10-14yrs old children and age group C to 15-18yrs old children.

3.3.3. The diversity and composition of core microbiota along the intestine of healthy children.

The core microbiota refer to the microbes that are shared amongst all, or the vast majority, of individuals within a tested group [266]. These shared microbiota represent the bacteria conserved during the mutual coevolution of humans and their intestinal microbes; together, these are relevant to human health [266]. Therefore, we characterized the core microbiota at three different intestinal regions (LC, RC and TI) of healthy children. The core microbiota at LC or RC of non-IBD children were smaller than the core microbes at TI (Figure 3.3.5 A). For statistical comparison of the diversity and the composition of the core microbiota at different sites, we identified the core microbiota that are shared amongst 75% of samples within each group. Although the core microbiota of the TI exhibited a significant higher richness as compared to LC or RC microbiota (Figure 3.3.5 B), no significant difference in diversity was detected amongst the three core microbiota (Figure 3.3.5 C). A Kruskal-Wallis test followed by Dunn's test for multiple comparisons identified 109 differentially abundant core OTUs among the three tested sites (Appendix VII). The LC core microbiota lacked OTUs belonging to *Atopobium*, *Carnobacterium*, *Vagococcus*, *Streptococcus pseudoneumonia*, *Blautia producta*, *Peptostreptococcus*, *E. rectale* and *Ferrimonas* while the OTUs assigned to *Adlercreutzia*, *Mogibacterium*, *Oscillospira*, and Ruminococcaceae were not identified in the RC and TI core microbiota (Appendix VII). While the OTUs assigned to *Xanthomonas*, *Pseudomonas*, *Subdoligranulum*, Alkaligenaceae, *Lactobacillus intestinalis*, *Anaerostipes* and *Blautia* were specific to the RC core microbes, the OTUs related to *Actinomyces oris*, *Parabacteroides distasonis* and *Oscillospira* were not found in the RC core microbes (Appendix VII). Finally, the TI core microbiota were characterized by an enrichment of OTUs belonging to *Blautia*, *Butyrivibrio*,

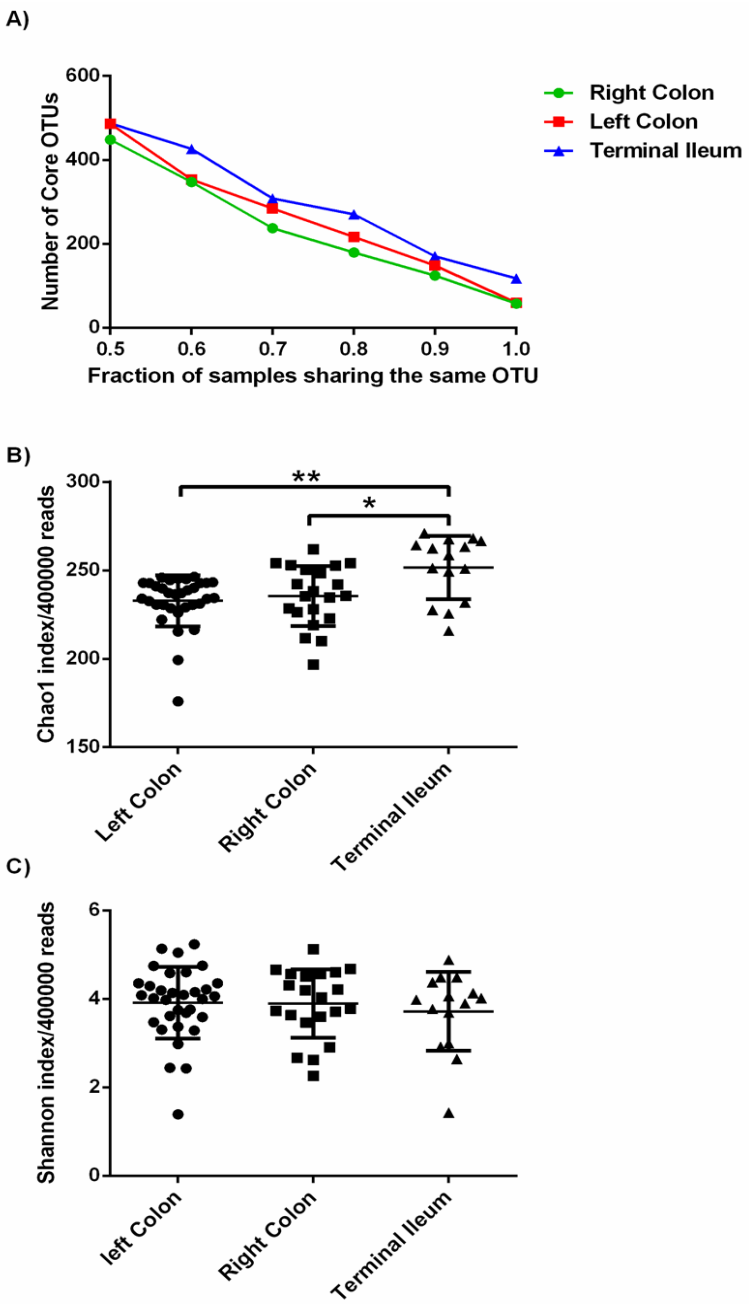


Figure 3.3.5: The left and right colon of non-IBD children harbour a less diverse core microbiome as compared to the terminal ileum. A, The size of the core microbiota shared by different sample percentages from right colon (green), left colon (red) and terminal ileum (blue). OTU corresponds to operating taxonomic unit. **B-C,** Chao1 and Shannon indices of the 0.75 core microbiota, respectively. Sequences assigned to the OTUs that are shared between 75% samples from each category (left colon (n=34); right colon (n=21) and terminal ileum (n=15)) were extracted and the diversity indices were calculated per 400,000 reads from each sample. The horizontal lines represent mean±SD. A Kruskal-Wallis test followed by Dunn's multiple comparisons was applied for statistical analysis (*p< 0.05 and ** p<0.01).

Oribacterium sinus, *Lachnospira*, *Faecalibacterium*, *Leclercia*, *Nitrincola*, and *Enterobacter hormaechei* and the depletion of 8 Lachnospiraceae OTUs (Appendix VII). Together, these observations indicate that the diversity and the composition of the core microbiota vary along the intestinal tract of non-IBD children.

3.4. Characterization of left colon (LC) microbiota in pediatric IBD.

3.4.1. Microbial diversity of the left colon in pediatric IBD.

The diversity of the gut microbiota has been reported to decrease in adults with IBD as compared to non-IBD controls [94]. However, the decrease in microbial diversity in pediatric IBD is still controversial [139, 140]. In order to assess the diversity of the left colon (LC) microbiota in first onset pediatric IBD, we sequenced the V6 hypervariable region of the 16S rRNA gene using mucosal aspirates as microbiota sampling method. A total of 80,059,639 high quality reads were obtained from 86 LC mucosal aspirates representing 34 non-IBD controls, 39 CD and 13 UC subjects. Closed reference OTU picking identified 5998 bacterial OTUs that were reduced to 3250 OTUs by filtering out singletons. Good's coverage score was above 99.9% for each sample. To determine the richness and diversity of the LC microbiota, we calculated the Chao1 and Shannon indices using randomly selected 499,620 reads from each sample. No significant difference in either the richness or the diversity of the LC microbiota was detected by comparing either CD or UC to controls (Figure 3.4.1 A-B). To test the association between disease location and the diversity of the LC microbiota, CD subjects were categorized according to the biogeographical location of the inflammation. The diversity of the LC microbiota was identical between ileal CD, colonic CD, ileal-colonic CD and control individuals (Figure 3.4.1 C-D).

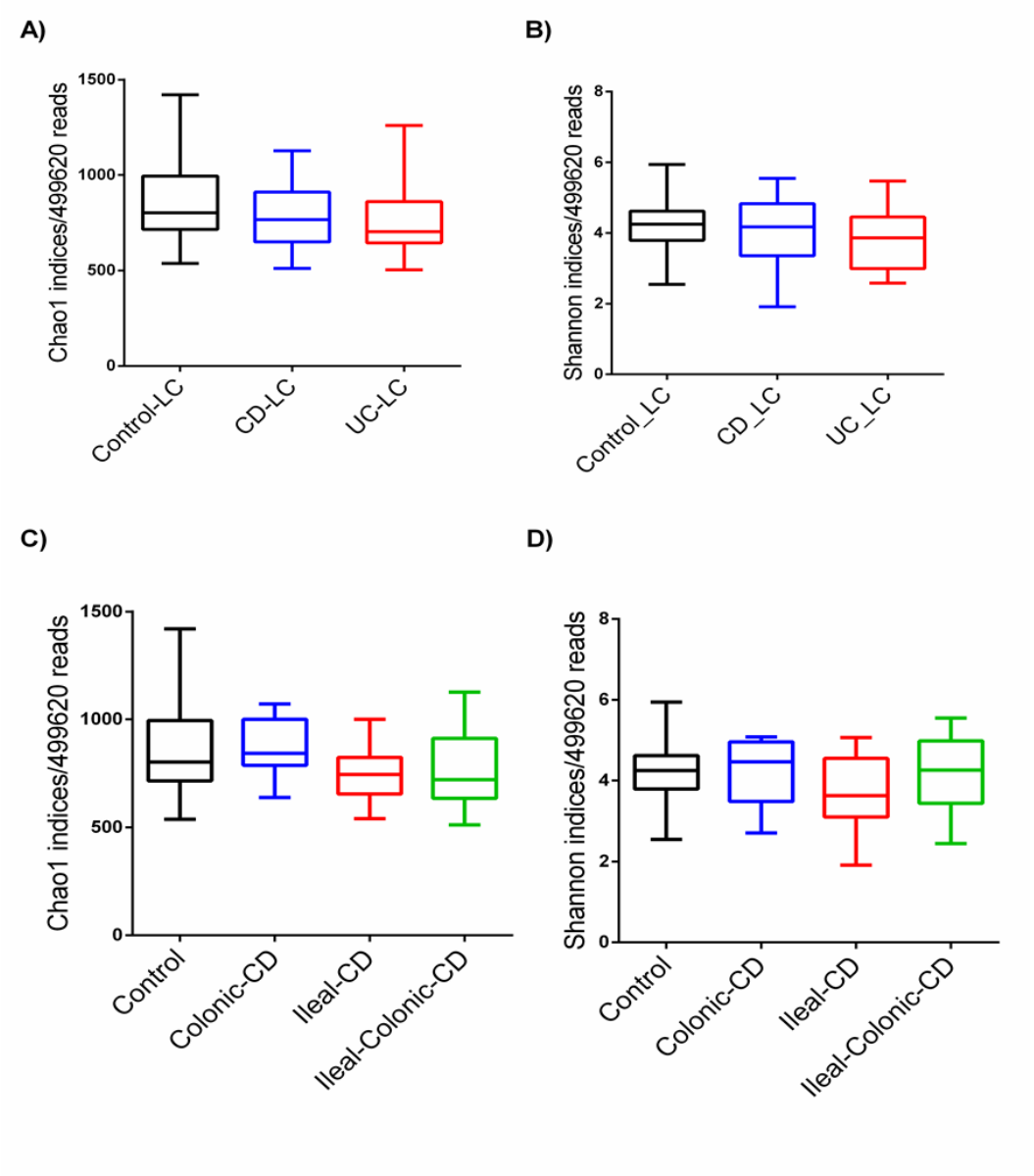


Figure 3.4.1: The gut microbiota exhibit a similar diversity at the left colon of pediatric patients with IBD as compared to control children. Richness and diversity of gut microbiota were estimated via Chao1 and Shannon indices calculated by QIIME 1.7. **A-B,** Diversity measures of RC microbiota in patients with Crohn's disease (CD; n= 36), ulcerative colitis (UC; n=12) and control subjects (n= 31) based on analyses of the Illumina reads. **C-D,** Diversity measures of RC microbiota in CD samples categorized by the disease location (Colonic CD, n= 9; Ileal CD, n= 12 and Ileal-colonic, n= 15). A Kruskal-Wallis test followed by Dunn's multiple comparisons were applied for statistical comparison.

3.4.2. Dysbiosis of the left colon (LC) microbiota in pediatric IBD.

Overall, a total of 16 bacterial phyla were identified in control, CD and UC groups. Actinobacteria, Bacteroidetes, Firmicutes, Fusobacteria, Proteobacteria and Tenericutes represented > 99% of all 16S rRNA sequences (Figure 3.4.2 A). Compared to controls, the relative abundance of Tenericutes was significantly lower in CD and UC ($p< 0.05$), while the relative abundance of Actinobacteria was significantly higher in UC only ($p< 0.01$, Figure 3.4.2 B). To identify LC microbes associated with IBD, we compared the microbial taxa composition using a non-parametric Kruskal-Wallis test, which identified 35 differentially abundant taxa (Bonferroni corrected $p< 0.0167$; Appendix VIII). Linear discrimination analysis (LDA) on the relative abundance of all identified taxa showed significant discriminating factors in 20 taxa between control microbiota and each of the CD and UC microbiota (Figure 3.4.3). While Tenericutes, *Clostridium*, *Eubacterium rectale*, *Ruminococcus*, *Lachnospira* and *Paenibacillus* were decreased in CD as compared to control microbiota, *Faecalibacterium*, Ruminococcaceae, Veillonellaceae, Gemellaceae, *Neisseria* and *Gemella* were enriched (Figure 3.4.3 A and C). The UC microbiota, on the other hand, was characterized by increased Actinobacteria, *Veillonella*, *Neisseria*, *Vitreoscilla*, *Actinobacillus*, *Leptotrichia* and *Peptostreptococcus*, with depletion of Tenericutes, *Clostridium*, Porphyromonadaceae, *Ruminococcus torque*, *Subdoligranulum* and *Parabacteroides* as compared to the control microbiota (Figure 3.4.3 B and D). The relative abundance of differential OTUs among control, CD and/or UC were further analyzed by partial least squares discriminant analysis (PLS-DA). The samples from each group separated as a distinct cluster (Figure 3.4.3 E). An acceptable PLS-DA model was obtained with 3 components (predictive ability parameter [Q^2 cum] = 0.31, goodness-of-fit parameter

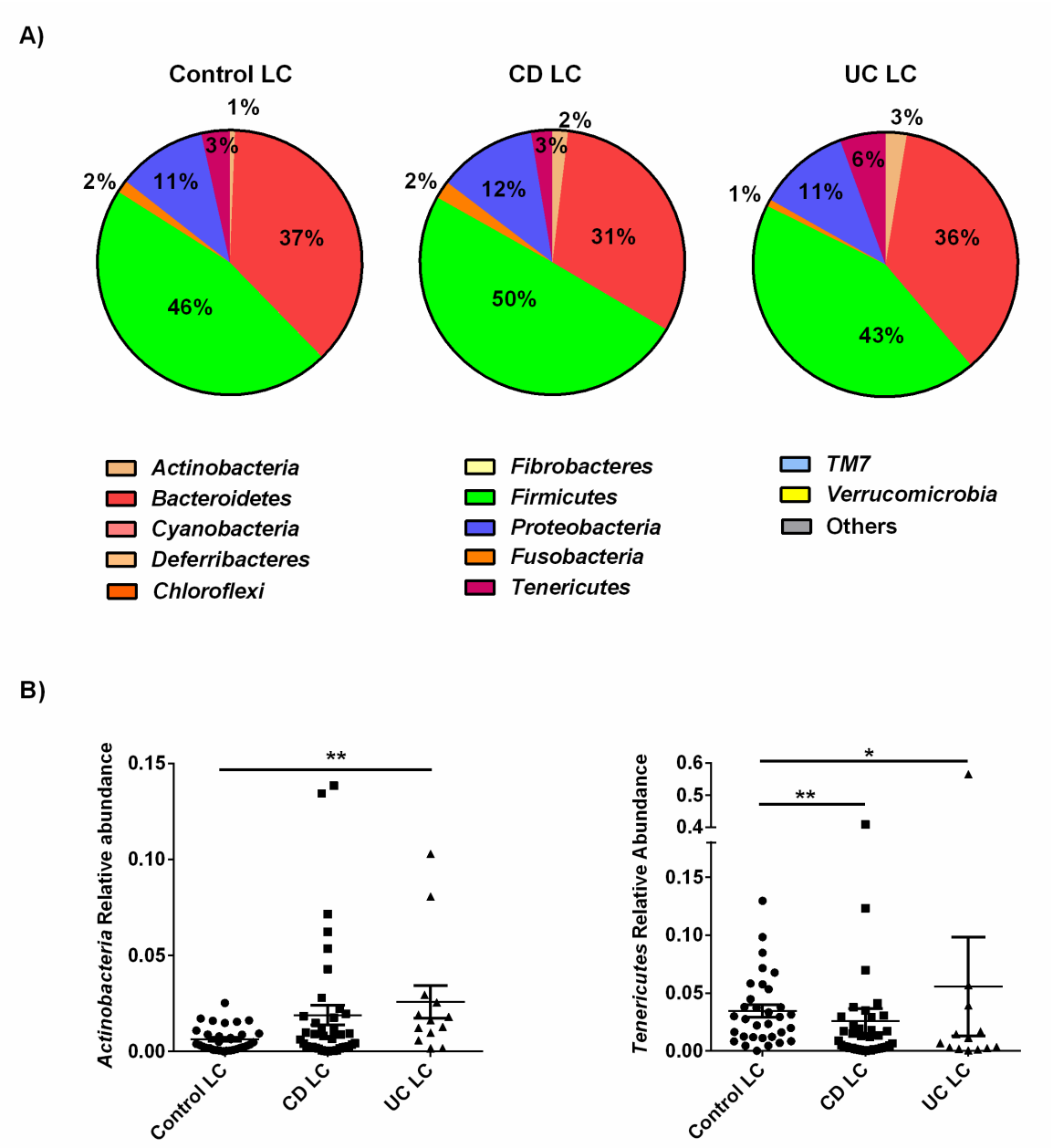


Figure 3.4.2: Taxonomic composition of the left colon microbial phyla of IBD patients and control subjects. A, Average relative abundance of bacterial phyla identified in patients with Crohn's disease (CD; n=39), ulcerative colitis (UC; n=13) and control subjects (n=34). **B,** Change in relative abundance of *Actinobacteria* and *Tenericutes* in controls, CD and UC patients. Kruskal-Wallis followed by Dunn's correction for multiple comparisons were applied for statistical comparison (* $p< 0.05$; ** $p< 0.01$). The horizontal line represents the mean relative abundance ±SEM in each group.

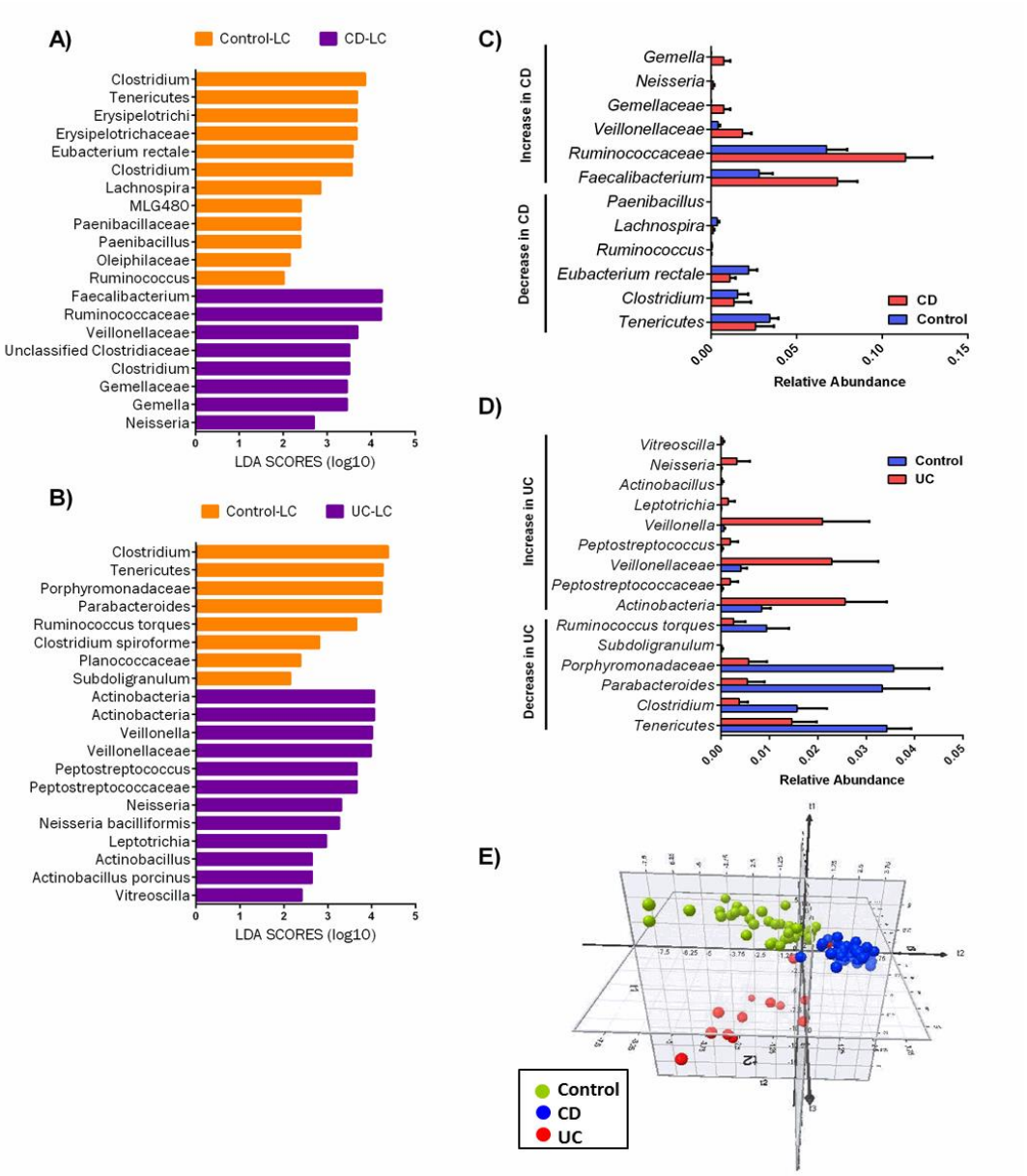


Figure 3.4.3: Dysbiosis of the left colon microbiota in pediatric IBD. The relative abundance of the bacterial taxa obtained from the analysis of the Illumina HiSeq reads were analysed by linear discrimination analysis (LDA) followed by effect size assessment using LEfSe [239]. **A,** Histogram of the LDA effect size score for CD-specific differentially abundant taxa (n=39 and 34 for CD and controls, respectively). **B,** Histogram of the LDA effect size score for UC-specific differentially abundant taxa (n=34 for control and n=13 for UC). **C-D,** Relative abundance of the major taxa from panel A and B. E, Partial least squares analysis (PLS-DA) of left colon microbiota from control subjects (n=34) against CD (n=39) and UC patients (n=13). The model was constructed based on the differentially abundant OTUs as determined by Kruskal-Wallis test followed by Bonferroni correction for multiple comparisons. The PLS-DA model showed predictive ability parameter [Q^2 cum] = 0.31, goodness-of-fit parameter [R^2Y cum] = 0.752 and prediction accuracy of 96.51%.

[R^2Y cum] = 0.752 and prediction accuracy of 96.51%). These results reveal the dysbiosis of the LC microbiota in pediatric patients with IBD.

3.4.3. Depletion of the LC core microbiota in pediatric patients with IBD.

The large microbial variabilities amongst individuals hinder the identification of disease specific biomarkers [266]. Therefore, more interest should be given to the common core microbiota that is conserved within each disease phenotype or the control group. Here, we defined the core microbiota as the microbes that are shared among 75% of individuals within each group. The LC core microbiota in CD and UC were smaller than the core microbiota in healthy children. A total of 281 OTUs (260 control, 189 CD and 226 UC) were shared by 75% of the samples within each group (Figures 3.4.4 and 3.4.5 A). A Kruskal-Wallis testing identified 52 core OTUs exhibiting differential abundance between at least two of the three groups ($p< 0.015$; Appendix IX). Thirty two OTUs were related to Lachnospiraceae which exhibited an increased abundance in the control core microbiota as compared to CD and/or UC. Compared to the control core microbiota, CD was characterized by a higher abundance of *Fusobacterium* and *Gemella* in association with a decreased abundance of the Tenericutes genus *MLG480* and 3 OTUs related to *Bacteroides* (Appendix IX and figure 3.4.5 B). The core microbiota of UC harbour more *Veillonella* and *Leptotrichia* with less *Alistipes* in comparison with the control group (Appendix IX and figure 3.4.5 B). In addition to the altered taxonomic composition of core microbiota in pediatric IBD, the microbial richness, but not the diversity (as expressed by Chao1 and Shannon indices, respectively), was significantly lower in the LC core microbiota of CD and UC subjects as compared to controls. Furthermore, the richness of the LC core microbiota from CD subjects was significantly lower than from UC patients (Figure 3.4.5 C-D).

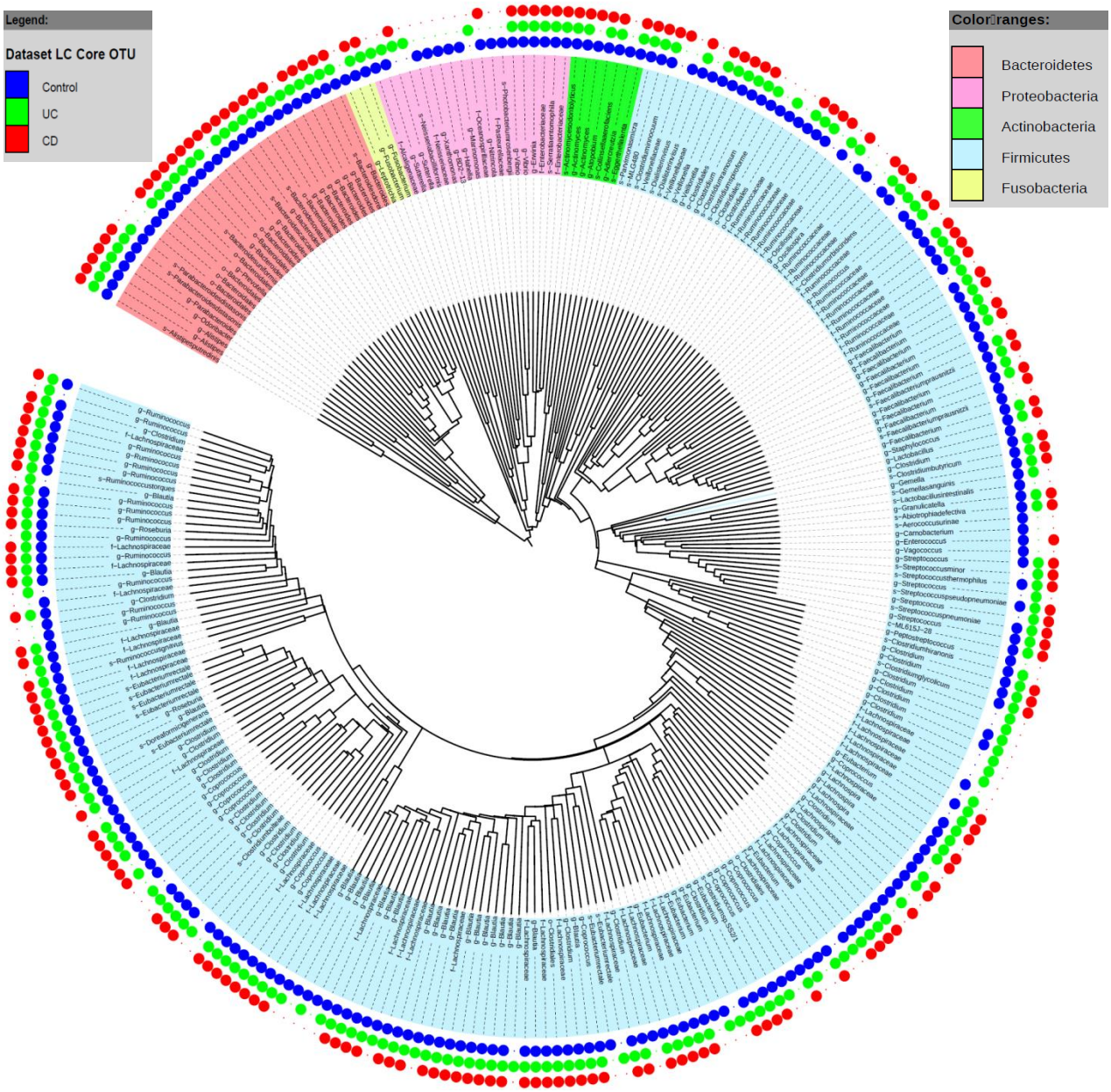


Figure 3.4.4: Phylogenetic tree of the microbial taxa detected in the left colon of at least 75% of the samples within each patient cohort. The figure was generated using the iTOL (Interactive Tree of Life) web package (http://itol.embl.de/upload.cgi; [243]). Different phyla are denoted in different colours. Taxa marked with a blue circle were identified as members of the core microbiota of the control subjects. Taxa marked with green or red circles were identified as members of the core microbiota of the UC or CD patients, respectively.

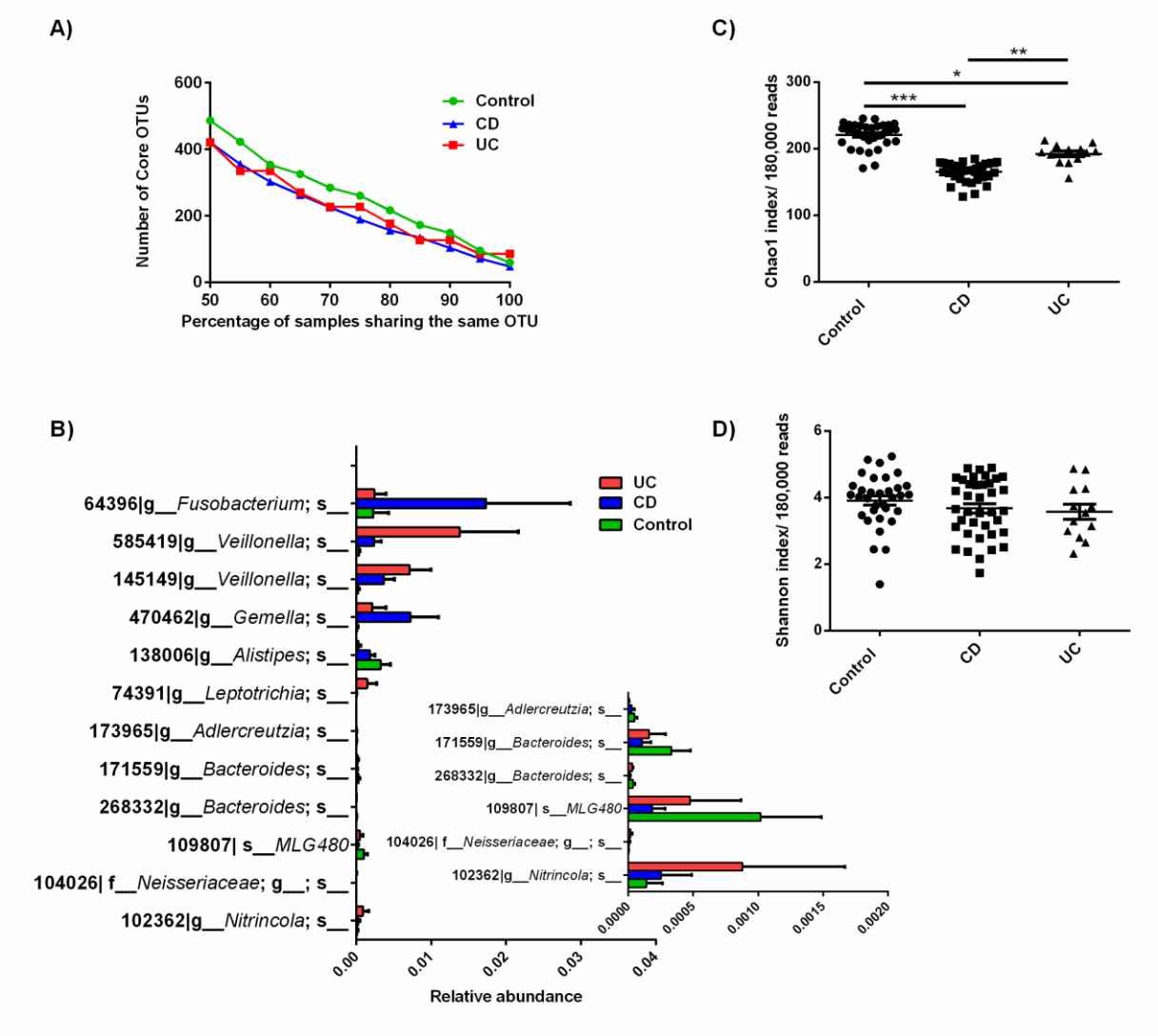


Figure 3.4.5: CD and UC microbial communities are characterized by a smaller and specific left colon core microbiome as compared to controls. A, The size of the core microbiota shared by different sample percentages of control subjects (green), CD (blue) and UC (red) patients. OTU corresponds to operating taxonomic unit. **B,** Relative abundance of the differentially abundant core OTUs among CD, UC and control groups as revealed by Kruskal-Wallis followed by Dunn's post hoc test. The inset figure shows the relative abundance of taxa present at low abundance. The y axis shows the OTU number and its corresponding taxonomy. **C** and **D,** Chao1 and Shannon indices of 0.75 LC core microbiota, respectively. Sequences assigned to the OTUs that are shared between 75% samples within each category (Control (n=33); CD (n=38) and UC (n=13)) were extracted and the diversity indices were calculated per 180,000 reads from each sample. The horizontal lines represent mean±SEM. Kruskal-Wallis test followed by Dunn's multiple comparisons were applied for statistical analysis (*p< 0.05; ** p<0.01 and *** p<0.001).

3.4.4. The left colon microbiota is associated with IBD activity.

In order to identify bacterial taxa that may contribute to the disease activity, we tested the association between the relative abundance of the identified taxa at different taxonomic levels and the severity of the inflammation as indicated by the disease activity index. A Kruskal-Wallis test identified 18 unique taxa exhibiting differential abundance between CD patients with mild, moderate or severe inflammation (Appendix X). Partial least squares discriminant analysis (PLS-DA) clustered CD patients with severe inflammation separately from those with mild or moderate inflammation (Figure 3.4.6 A). Taxa biplot analysis indicated that certain taxa, namely *Allobaculum SPID4, Oribacterium sinus, Selenemonas sputigena, Pantoea,* Clostridiales FamilyXI IncertaeSedis, *Streptococcus pneumonia, Atopobium* and *Atopobium parvulum,* were more abundant in patients with severe, as compared to mild, inflammation (Figure 3.4.6 C). In contrast, Clostridia and *Facklamia* were enriched in samples from mild inflammation relative to the severe inflammation and exhibited a significant decline in relative abundance with increased disease severity. On the other hand, a Kruskal-Wallis test identified 28 differentially abundant taxa between UC patients with mild, moderate and/or severe inflammation (Appendix XI). Partial least squares discriminant analysis (PLS-DA) generated 3 separate clusters of UC samples according to the disease activity (Figure 3.4.6 B). Taxa biplot correlation analysis indicated the higher relative abundance of *Fibrobacter intestinalis, Staphylococcus aureus, Macrococcus caseolyticus, Enterococcus haemoperoxidus, Abiotrophia defective, Neisseria bacilliformis, Corynebacterium durum* and their genera in the microbiota of UC patients with severe, as compared to mild, inflammation. In contrast, *Lachnospira,* Lachnospiraceae, Bacteroidia, *Eubacterium, Coprococcus, C. hylemonae* and *Dialister pneumosintes* were more abundant in UC patients with mild, as compared to severe, inflammation (Figure 3.4.6 D).

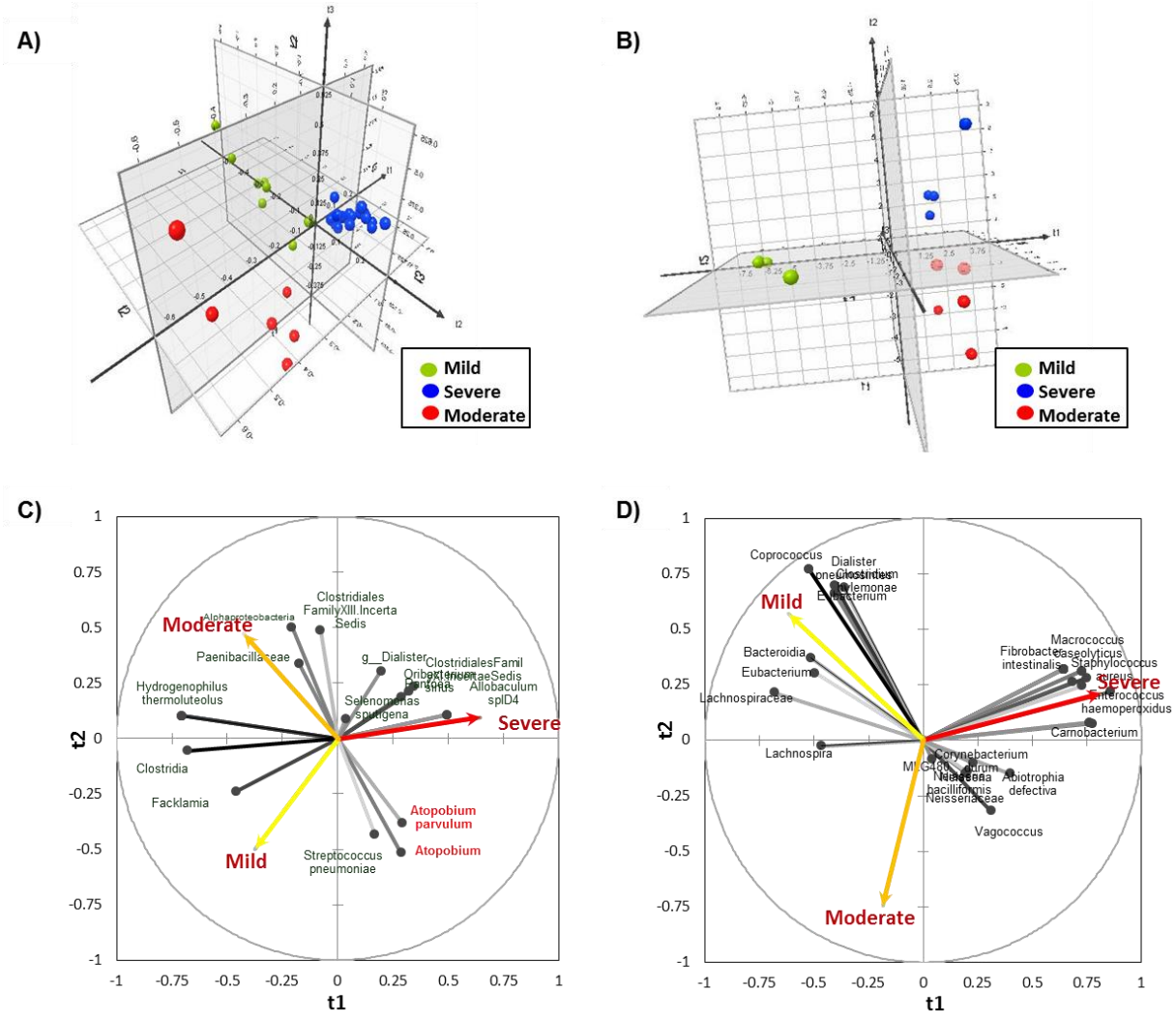


Figure 3.4.6: Partial least squares analysis of the gut microbiota revealed biomarkers of disease severity. A, PLS-DA of CD patients with severe inflammation (n=23) against CD patients with mild inflammation (n=10) and CD patients with moderate inflammation (n=6). B, PLS-DA of UC patients with severe inflammation (n=4) against CD patients with mild inflammation (n=4) and CD patients with moderate inflammation (n=5). The model was constructed based on the differentially abundant taxa as determined by Kruskal-Wallis test followed by Bonferroni correction for multiple comparisons. The PLS-DA model showed the predictive ability parameter [Q^2 cum] = 0.379 and 0.643, goodness-of-fit parameter [R^2Y cum] = 0.792 and 0.903 and prediction accuracy of 92.31% and 100% for CD and UC, respectively. C-D, Biplot analysis of the first two components of the PLS-DA models in panel A and B showing the significant taxa relative to disease activity (arrows).

3.5. Characterization of the right colon (RC) microbiota in pediatric IBD.

3.5.1. Microbial diversity of the right colon in pediatric IBD.

Eighty to ninety percent of children with either CD or UC suffer from RC inflammation [4, 5]. Therefore, we aimed to characterize the diversity and the composition of the RC microbiota in first onset pediatric IBD. The microbial diversity was characterized in RC mucosal aspirates from 72 subjects (21 controls, 37 CD and 14 UC) by sequencing of the V6 regions of the 16S rRNA gene using Illumina HiSeq 2500 and confirmed on a subset of samples (9 Controls, 9 CD and 8 UC) using 454-pyrosequencing. Illumina sequencing generated 71,992,878 high-quality reads with an average of 999,901.08±332,290 reads per sample. Closed reference OTU picking assigned these reads to 5998 OTUs (0.97), which were reduced to 3317 OTUs after filtering out the singletons with a Good's coverage of > 99.9% for each sample. RC microbiota richness and diversity amongst the three groups (Control, CD and UC) were compared by estimating Chao1 and Shannon indices. No significant difference in richness or diversity of the RC microbiota was observed in either CD or UC compared to the control group (Figure 3.5.1 A-B). Moreover, no differential diversity was detected among control and CD patients with either ileal, colonic or ileocolonic inflammation (Figure 3.5.1 C-D). These results were confirmed by 454-pyrosequencing which did not detect any difference in the richness or diversity between any pairwise comparison of the three different groups (Control, CD and UC; Figure 3.5.1 E-F).

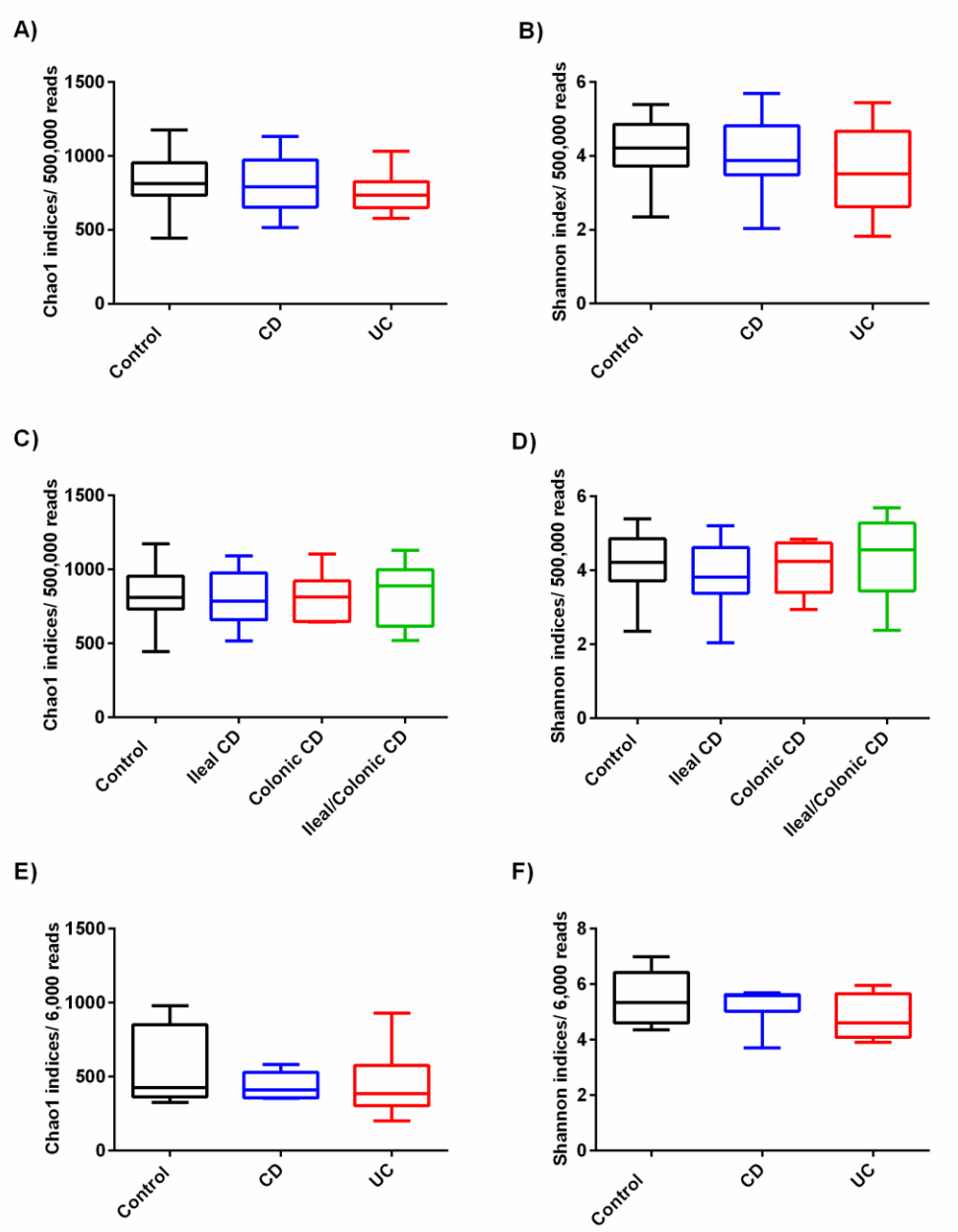


Figure 3.5.1: The gut microbiota exhibit similar diversity at the right colon of pediatric patients with IBD in comparison to control children based on Illumina HiSeq 2500 and 454-pyrosequencing. Richness and diversity of the gut microbiota were estimated via Chao1 and Shannon indices as calculated by QIIME 1.7. **A-B,** Diversity measures of RC microbiota in patients with Crohn's disease (CD; n= 37), ulcerative colitis (UC; n=14) and control subjects (n= 21) based on Illumina HiSeq 2500 analyses. **C-D,** Diversity measures of RC microbiota in CD samples categorized by the inflamed intestinal site (Colonic CD, n= 6; Ileal CD, n= 16 and Ileal-colonic, n= 14). **E-F,** Diversity indices of RC microbiota in the three main cohorts (CD, n= 9; UC, n=8 and control, n=9) based on 454-pyrosequencing results. A Kruskal-Wallis test followed by Dunn's multiple comparisons were applied for statistical comparison.

3.5.2. Dysbiosis of the RC microbiota in pediatric IBD.

The RC microbiota of CD patients harbored a total of 18 phyla while the UC and control microbiota included 16 phyla (Figure 3.5.2 A). Lentisphaerae was identified only in CD microbiota. In addition, Chloroflexi and Synergistetes were missing in the microbiota of healthy and UC children, respectively. More than 99% of the 16S rRNA reads were assigned to only 6 phyla; Actinobacteria, Bacteroidetes, Fusobacteria, Firmicutes, Proteobacteria and Tenericutes. Compared to the control microbiota, Firmicutes and Verrucomicrobia showed lower relative abundance in children with UC (56.35±20.15% vs 35.95±5.02% for Firmicutes and 0.017±0.010% vs $1.941x10^{-4}$±$1.428x10^{-4}$% for Verrucomicrobia; Figure 3.5.2 B). On the other hand, no phyla exhibited differential abundance between the RC microbiota of CD and control groups. At lower phylogenetic levels, a Kruskal-Wallis test indicated significant discriminating factors in 48 taxa (p<0.015; Appendix XII). The relative abundance of Clostridia, Clostridiales, and Lachnospiraceae, were decreased in CD and UC as compared to control microbiota while Negativicutes, *Selenomonadales*, *Veillonella* and Betaproteobacteria were increased (Appendix XII). Linear discriminant effect size analysis (LefSe) found decreased relative abundance of Lachnospiraceae (30.19±3.5% in CD vs 44.4±4.4% in control), Clostridiaceae (1.54±1.1% in CD vs 2.54±1.1 in control) and Erysipelotrichaceae (3.02±1.3% in CD vs 3.65±1.03% in control) in association with increased abundance of *Faecalibacterium* (5.36±1.1% in CD vs 2.43±1.09% in control), Betaproteobacteria (2.88±0.86% in CD vs 0.83±0.45% in control), Veillonellaceae (1.75±0.5% in CD vs 0.3±0.15% in control), *Eggerthella lenta* (0.48±0.17 in CD vs 0.08±0.02% in control), *Haemophilus* ($6.2x10^{-5}$±$3.4x10^{-5}$% in CD vs $2.6x10^{-6}$±$1.3x10^{-6}$% in control), and *Bacteroides dorei* (4.5±1.4% in CD vs 2.5±1.4% in control) in CD patients as compared to controls (Figure 3.5.3). Dysbiosis of the microbiota of UC patients was

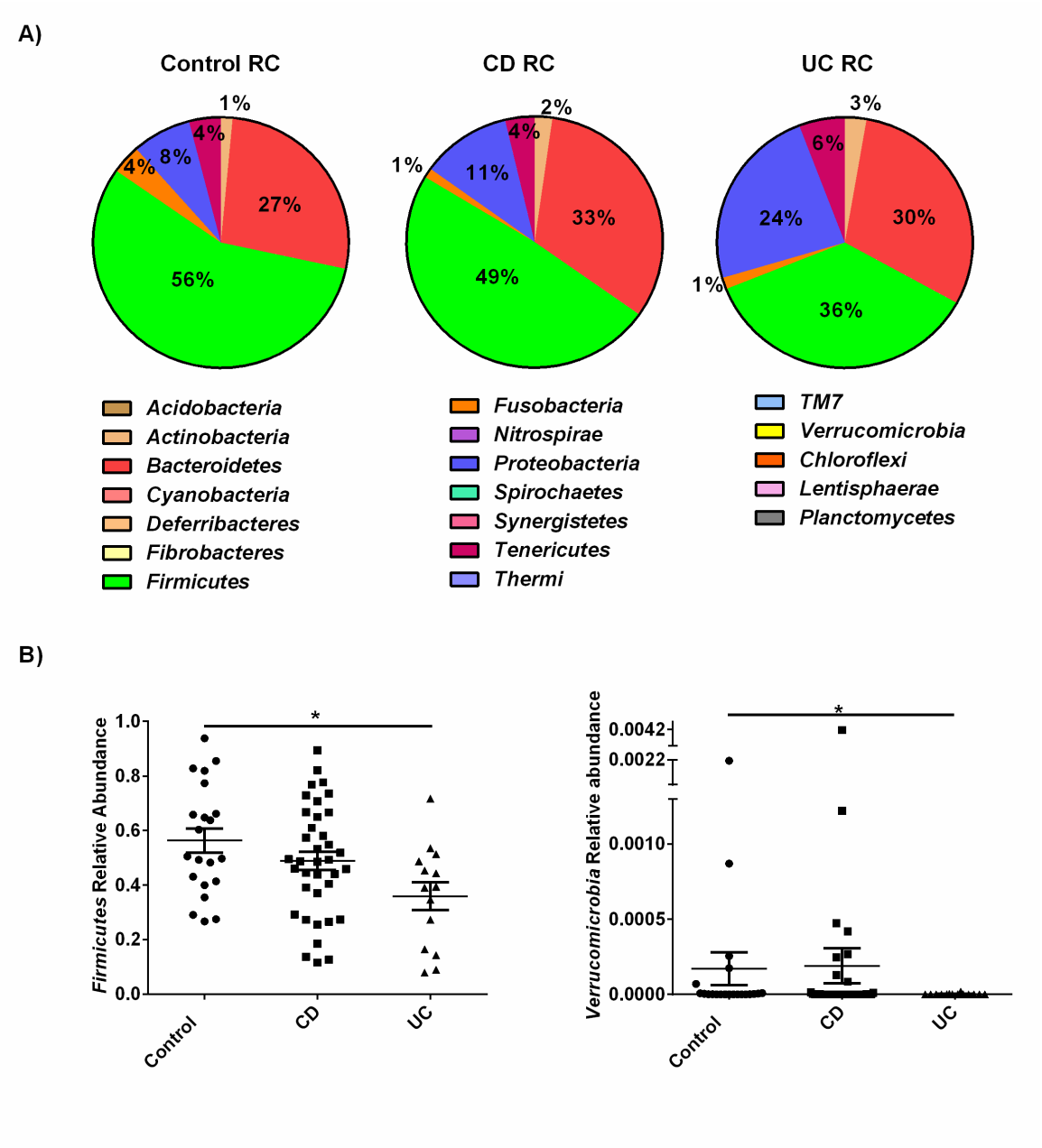


Figure 3.5.2: Taxonomic composition of the right colon microbiota of IBD patients and control subjects. A, Average relative abundance of bacterial phyla identified in patients with Crohn's disease (CD; n=37), ulcerative colitis (UC; n=14) and control subjects (n=21). **B,** Change in relative abundance of *Firmicutes* and *Verrucomicrobia* in controls, CD and UC patients. A Kruskal-Wallis followed by Dunn's correction for multiple comparisons test was applied for statistical comparison (*p< 0.05). The horizontal line represents the mean relative abundance ±SEM in each group.

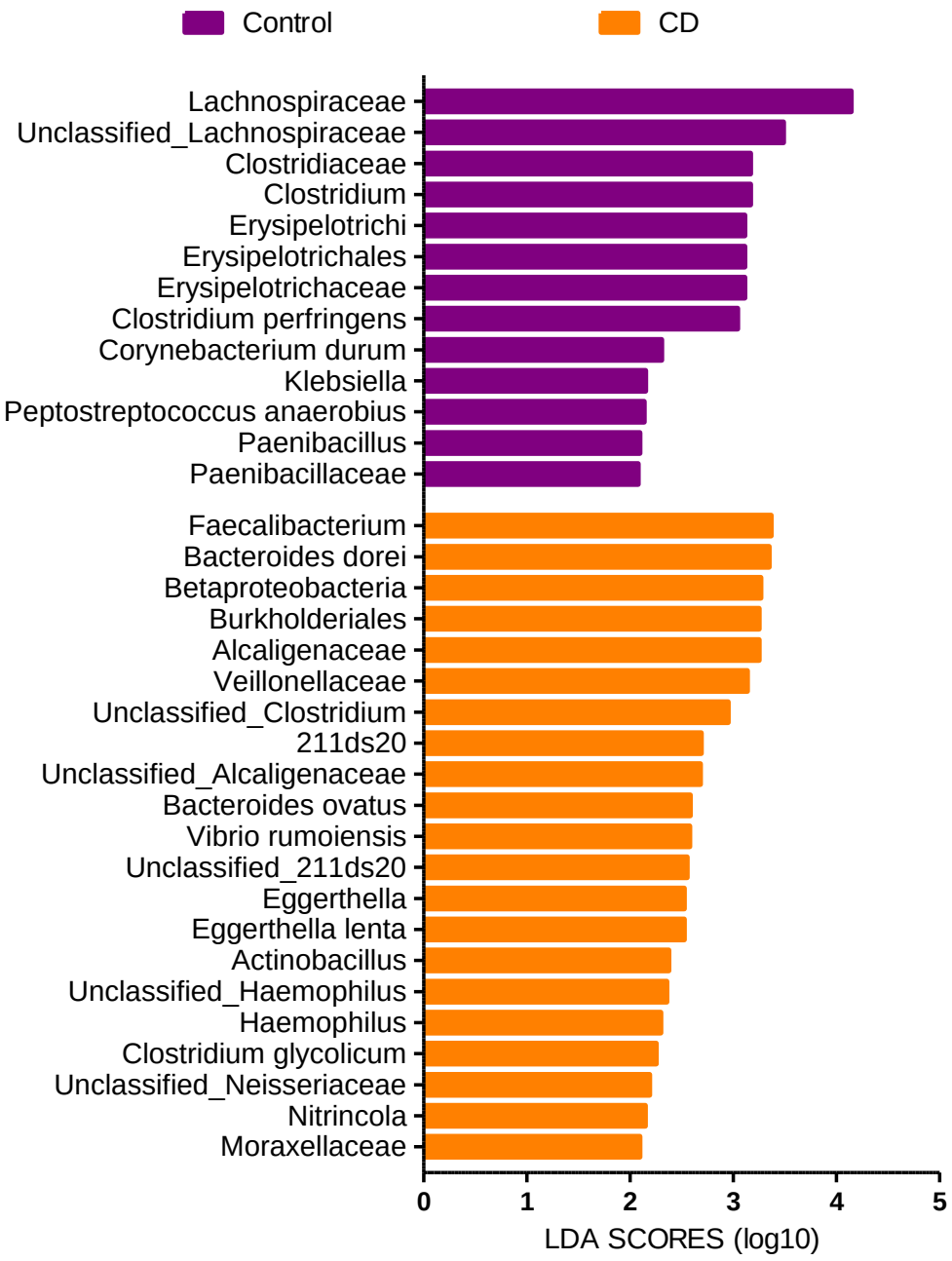


Figure 3.5.3: CD-specific microbial biomarkers identified at the right colon. Histogram of the LDA effect size score for CD-specific differentially abundant taxa (n=37 and 21 for CD and controls, respectively). The relative abundance of the bacterial taxa obtained from the analysis of the Illumina HiSeq reads were analysed by linear discriminant analysis (LDA) followed by effect size assessment using the LEfSe tool [239].

characterized by depletion of Firmicutes (56.35±20.15% vs 35.95±5.02%), Clostridiales (34.7±5% vs 54.9±4.2%), Lachnospiraceae (21.6±4% vs 44.4±4.4%), *Roseburia* (1.39±0.39% vs 5.31±1.3%), *Blautia* (2.56±0.6% vs 5.14±0.7%), *Ruminococcus* (0.01±0.005% vs 0.08±0.02%), *Coprococcus* (0.74±0.25% vs 1.05±0.20%), Verrucomicrobia ($1.94x10^{-4}$±$1.42x10^{-4}$% vs 0.017±0.010%) and *Akkermansia* ($1.74x10^{-4}$±$1.41x10^{-4}$% vs 0.017±0.010%) in association with enrichment of *Porphymonas* ($1.17x10^{-3}$±$6.35x10^{-4}$% vs $3.9x10^{-5}$±$1.5x10^{-5}$%), Veillonellaceae (2.1±0.7% vs 0.3±0.15%), Alphaproteobacteria, ($5.2x10^{-3}$±$2.98x10^{-3}$% vs $1.96x10^{-4}$±$5.38x10^{-5}$%) and different members of Beta- and Gammaproteobacteria as compared to controls (Figure 3.5.4).This analysis was strengthened by PLS-DA of the relative abundance of these 48 taxa, which revealed a clear separation between the IBD and control groups (Figure 3.5.5 A). Permutation testing (1000 times) indicated that the observed separations were unlikely to be due to chance ($p< 0.002$). Altogether these results indicate a radically different microbial community structure in IBD patients as compared to controls. Differences between the microbiota of CD and UC patients were further assessed by PLS-DA analysis, which resulted in a detectable and significant separation between CD and UC communities (with a $p < 0.004$ as determined by 1000 permutations; Figure 3.5.5 B). The alteration of the RC microbial composition in IBD was confirmed by 454-pyrosequencing approach on a subset of samples (9 Controls, 9 CD and 8 UC). A total of 266,006 high quality reads were obtained for this subset of samples, after excluding low quality and short reads, with an average of 10,223.96 sequences per sample and a mean length of 169.58 bases including the primers. Proteobacteria was the only phyla that exhibited a differential relative abundance in either CD or UC microbiota as compared to the control group ($p< 0.5$; Figure 3.5.6 A-B). Linear discriminant analysis indicated a major discriminating role of the increased relative

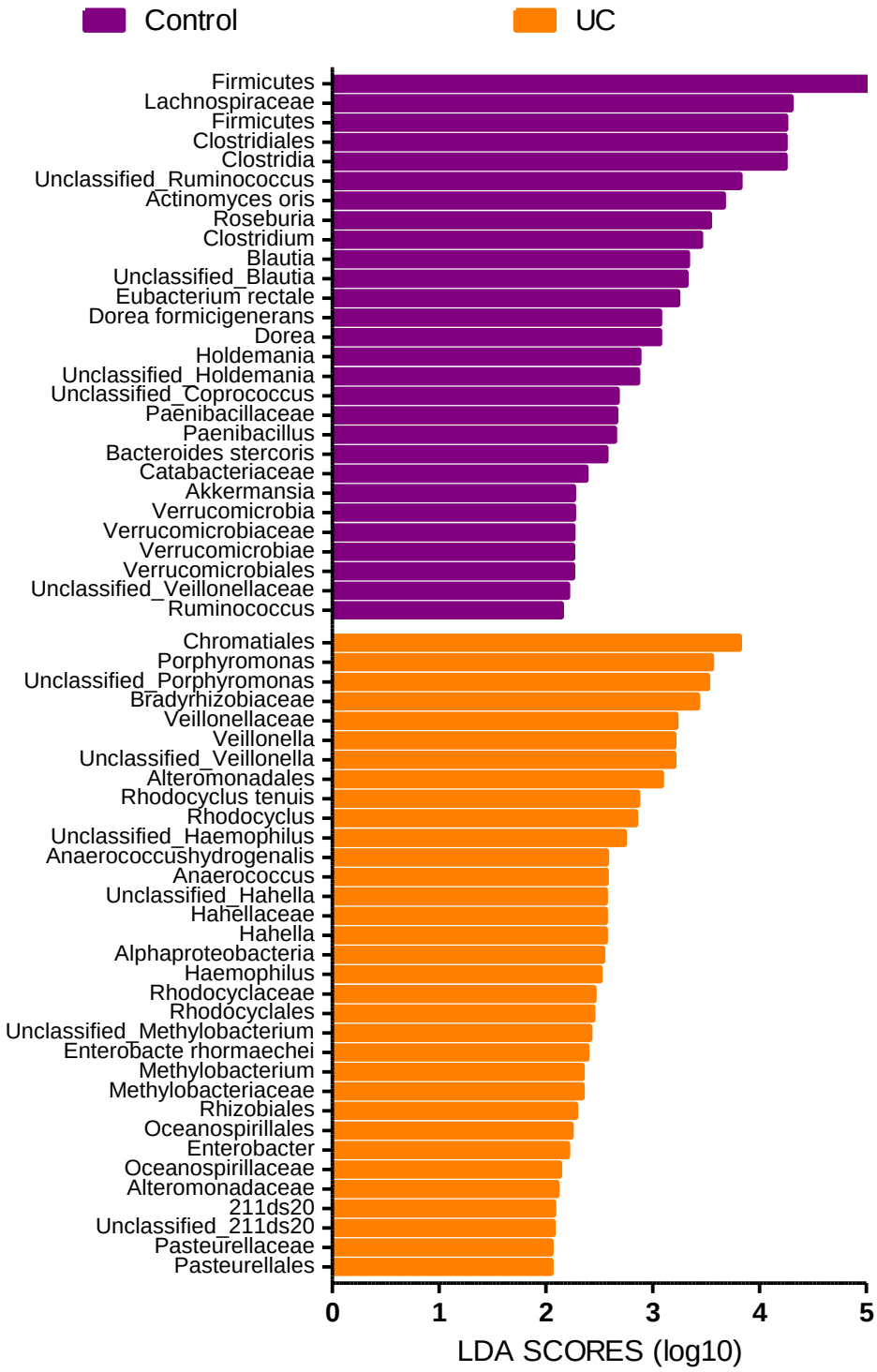


Figure 3.5.4: UC-specific microbial biomarkers identified at the right colon. Histogram of the LDA effect size score for UC-specific differentially abundant taxa (n=14 and 21 for UC and controls, respectively). The relative abundance of the bacterial taxa obtained from the analysis of the Illumina HiSeq reads were analysed by linear discriminant analysis (LDA) followed by effect size assessment using the LEfSe tool [239].

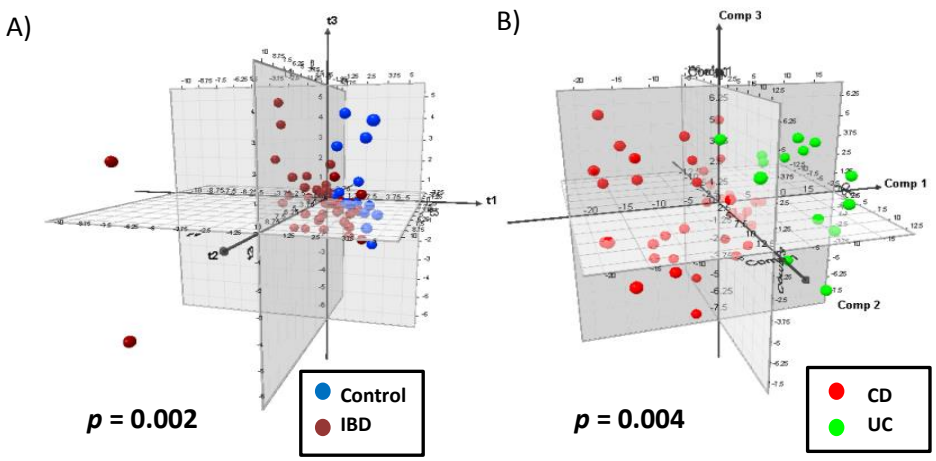


Figure 3.5.5: Dysbiosis of the right colon microbiota in pediatric IBD. A, PLS-DA of the right colon microbiota from IBD patients (n=51) against control subjects (n=21). **B,** PLS-DA of the right colon microbiota from CD patients (n=37) against UC patients (n=14). The model was constructed based on the differentially abundant OTUs determined by Kruskal-Wallis test followed by Bonferroni correction for multiple comparisons. To confirm the validation of the PLS-DA models, permutation tests (n=1000) were performed and the corresponding *p* value for prediction accuracy calculated.

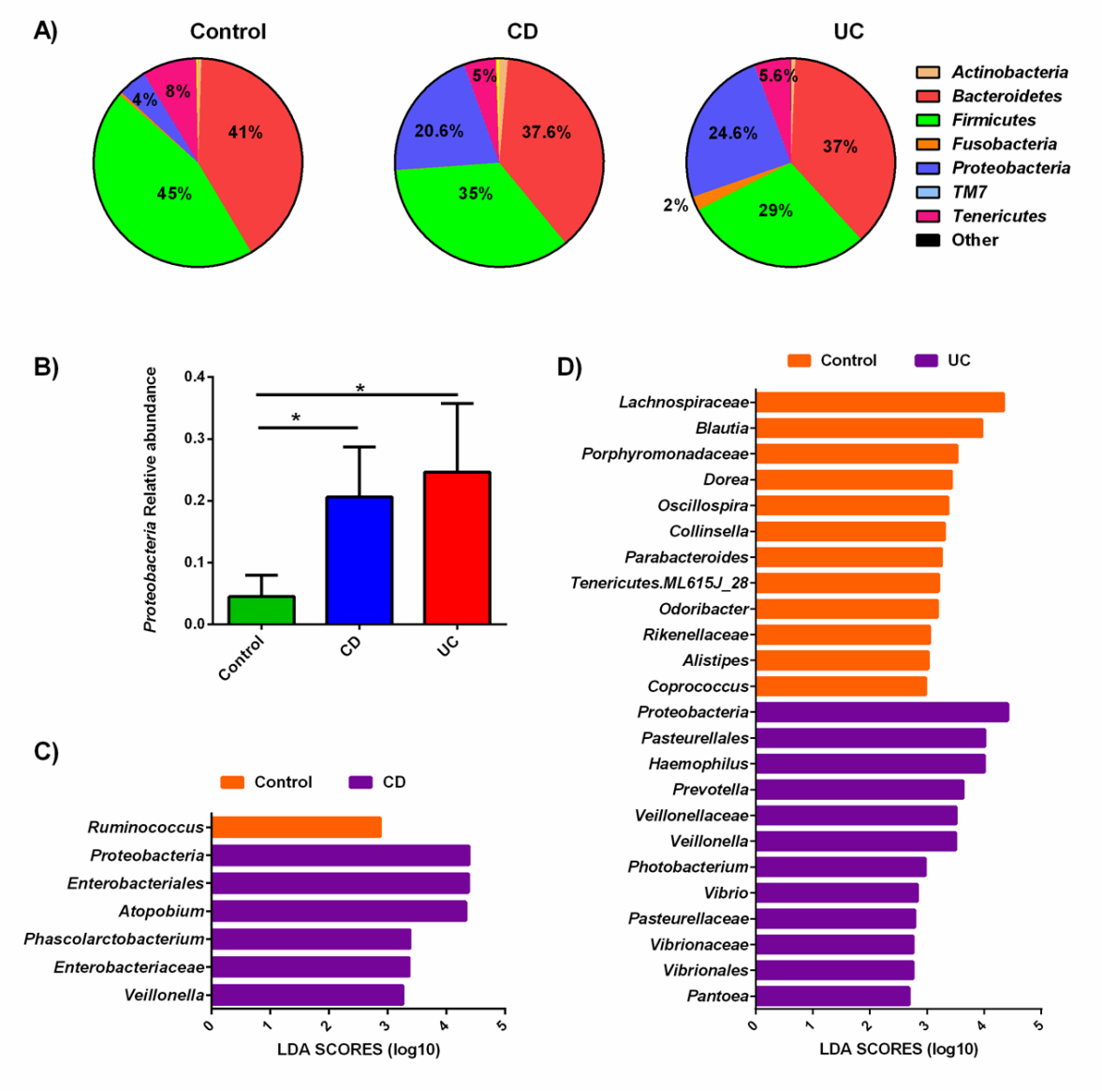


Figure 3.5.6: Taxonomic composition of the RC microbiota of IBD patients and control subjects based on 16S rRNA pyrosequencing reads. Metagenomic DNA extracted from 9 samples from non-IBD controls, 9 samples from CD and 8 samples from UC were subjected to 454-pyrosequencing and the generated raw data were analyzed using QIIME 1.4. **A,** Average relative abundance of bacterial phyla identified in patients with Crohn's disease (CD; n=9), ulcerative colitis (UC; n=8) and control subjects (n=9). **B,** Relative abundance of *Proteobacteria* (mean±SEM) in controls, CD and UC patients. A Kruskal Wallis test followed by Dunn's multiple comparisons were applied for statistical pairwise comparison (**p* <0.5). **C,** Histogram of the LDA effect size score for CD-specific differentially abundant taxa (n=9 for CD and controls). **D,** Histogram of the LDA effect size score for UC-specific differentially abundant taxa (n=9 for control and n=8 for UC).

abundance of Proteobacteria, *Atopobium*, *Phascolaractobacterium* and *Veillonella* and in the decreased abundance of *Ruminococcus* in dysbiosis of RC microbiota in CD (Figure 3.5.6 C). The microbiome of children with UC was also characterized by an enrichment of Proteobacteria members and *Veillonella*, and depletion of Lachnospiraceae, *Blautia*, *Dorea*, Porphyromondaceae, *Alistipes*, *Odoribacter*, *Oscillispira*, *Collinsella*, *Parabacteroides* and *Coprococcus* as compared to control microbiota (Figure 3.5.6 D).

3.5.3. Depletion of the RC core microbiota in pediatric IBD.

In order to reduce the impact of interindividual variabilities on the identified microbiota, we determined the microbiota that are conserved amongst 75% of the individuals within each test group. A total of 241 core OTUs were detected as RC core microbiota, which represent 90.2±8.3% of the total microbial population. CD and UC microbial communities were characterized by a smaller core microbiota as compared to controls with 179, 172 and 214 core OTUs for CD, UC and control subjects, respectively (Figures 3.5.7 and 3.5.8 A). A Kruskal-Wallis test identified 20 core OTUs exhibiting differential abundance between at least two of the three groups (CD, UC and controls; Appendix XIII). Only two taxa from the core microbiota exhibited increased abundance in CD or UC patients as compared to the control subjects. *Fusobacterium nucleatum* (OTU#64396) and *Veillonella parvula* were found to be more abundant in CD and UC patients, respectively. In contrast, Lachnospiraceae, *Fusobacterium nucleatum* (OTU#288565) and *Bacteroides* were the only taxa that exhibited a higher abundance in the control samples as compared to CD or UC groups (Figure 3.5.8 B and appendix XIII). A pairwise comparison between the core microbiota of CD and UC indicated that *Alistipes* was highly abundant in CD, while *Fusobacterium* (OTU#288565) showed an

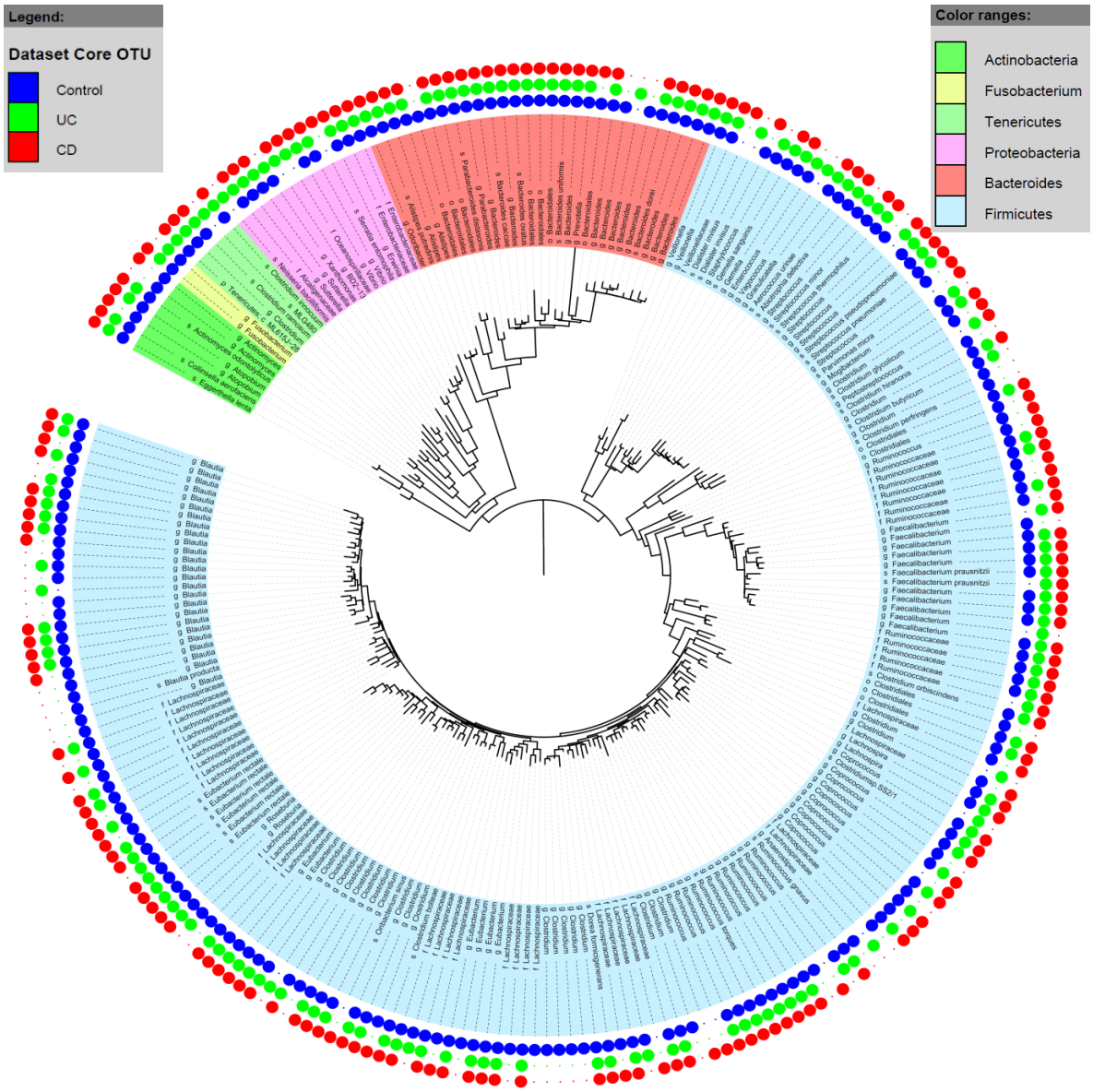


Figure 3.5.7: Phylogenetic tree of the microbial taxa detected in the right colon of at least 75% of the samples within each patient cohort. The figure was generated using the iTOL (Interactive Tree of Life) web package (http://itol.embl.de/upload.cgi; [243]). Different phyla are denoted with different colors. Taxa marked with a blue circle were identified as members of the core microbiota of the control subjects. Taxa marked with green or red circles were identified as members of the core microbiota of the UC and CD patients, respectively.

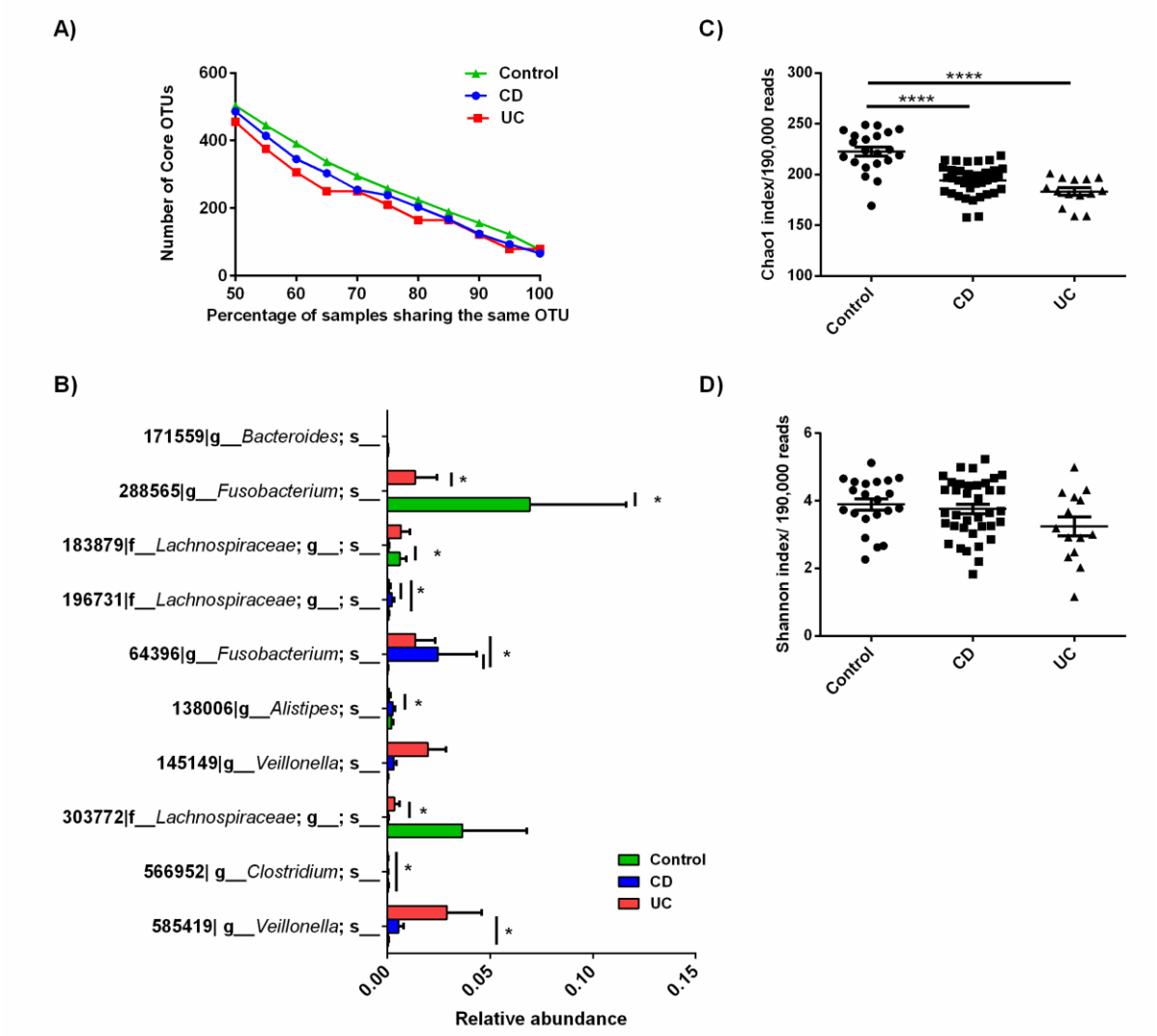


Figure 3.5.8: The CD and UC microbial communities are characterized by a smaller, specific right colon core microbiota as compared to controls. A, The size of the core microbiota shared by different sample percentages of control subjects (green), CD (blue) and UC (red) patients. OTU corresponds to operating taxonomic unit. **B,** Relative abundance of the statistically significant differential core OTUs among CD, UC and control groups as revealed by Kruskal-Wallis followed by Dunn's post hoc test. The y axis shows the OTU number and its corresponding taxonomy. **C** and **D,** Chao1 and Shannon indices of 0.75 RC core microbiota, respectively. Sequences assigned to the OTUs that are shared between 75% samples from each category (Control (n=21); CD (n=37) and UC (n=14)) were extracted and the diversity indices were calculated per 190,000 reads from each sample by QIIME. A Kruskal-Wallis test followed by Dunn's multiple comparisons were applied for statistical analysis (**** p <0.0001). The horizontal lines represent mean±SEM.

increased abundance in UC (Figure 3.5.8 B and appendix XIII). These microbial differences were reflected on the diversity of the core microbiota of different phenotypes with richness, but not eveness, significantly increased in the core microbiota of healthy subjects as compared to CD and UC groups (Figure 3.5.8 C-D).

3.5.4. The right colon microbial taxa are associated with pediatric IBD activity.

To explore whether disease activity can be identified by the RC microbial composition, a PLS-DA model was constructed to assess the stratification of CD and UC patients according to disease severity (Figure 3.5.9 A-B). First, a Kruskal-Wallis test identified 11 unique taxa exhibiting differential abundance between CD patients with mild, moderate and severe inflammation (Appendix XIV). Second, the significant taxa were used to construct the PLS-DA model. While PLS-DA could not reliably distinguish moderate from mild or severe inflammation, it clustered CD patients with severe inflammation separately from those with mild inflammation (Figure 3.5.9 A). This result indicates that the identified taxa could be used to stratify patients according to disease severity. Taxa biplot analysis indicated that the driving taxa are Clostridia and *Staphylococcus* for mild inflammation, *Phascolarctobacterium, Sutterella* and Betaproteobacteria for moderate inflammation, and Carnobacteriaceae, *Granulicatella, Mogibacterium, Proprionibacterium,* Bacillaceae and *Atopobium* for severe inflammation respectively (Figure 3.5.9 C). Phylotypes belonging to *Clostridia* exhibited a significant decline in relative abundance with increased disease severity. Reads assigned as *Atopobium* by QIIME analysis were retrieved and found to match to *A. parvulum* following read alignment against the RDB and NCBI databases (the aligned region covers the entire sequence length with >99% sequence identity to *A. parvulum*). On the other hand, a Kruskal-Wallis test identified 10 taxa with differential

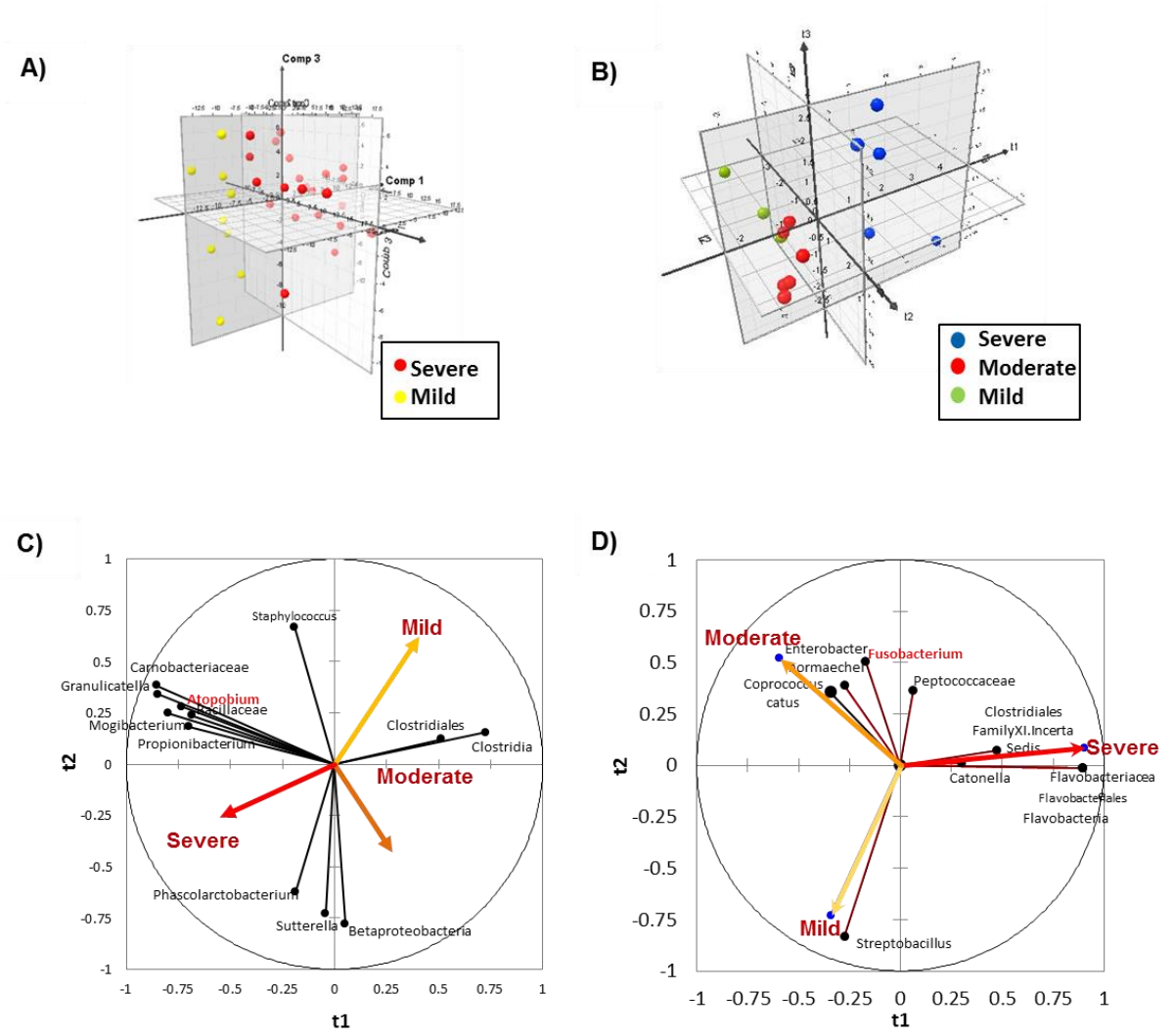


Figure 3.5.9: Partial least squares analysis of the right colon microbiota revealed biomarkers of disease severity. A, PLS-DA of CD patients with severe inflammation (n=23) against CD patients with mild inflammation (n=9). **B,** PLS-DA of UC patients with severe inflammation (n=5) against UC patients with mild inflammation (n=3) and UC patients with moderate inflammation (n=6). The model was constructed based on the differentially abundant taxa as determined by a Kruskal-Wallis test followed by Bonferroni correction for multiple comparisons. The PLS-DA model shows a predictive ability parameter [Q^2 cum] = 0.379 and 0.573, goodness-of-fit parameter [R^2Y cum] = 0.792 and 0.749 and prediction accuracy of 92.31% and 92.86% for CD and UC, respectively. **C-D**, Biplot analysis of the first two components of the PLS-DA models in panel A and B showing the significant taxa relative to disease activity (arrows).

relative abundance between UC samples with mild, moderate and/or severe inflammation (Appendix XV). PLS-DA analysis of the relative abundance of these different taxa indicated a clear separation between mild, moderate and severe UC samples (Figure 3.5.9 B). A taxa biplot correlated *Coprococcus catus* and *Enterobacter hormaechei* with moderate UC inflammation, *Catonella,* Flavobacteria, Flavobacteriales, Flavobacteriaceae and Clostidiales-FamilyXI.incertae Sedis with severe UC and *Streptobacillus* with mild inflammation. Two taxa, namely *Fusobacterium* and Peptococcaceae, were correlated with both moderate and severe inflammation (Figure 3.5.9 D).

3.5.5. qPCR validation of the Illumina sequencing results from RC samples.

To validate the results obtained by the Illumina sequencing approach, the relative abundance of the discriminant taxa, *Atopobium parvulum*, *F. nucleatum*, *Veillonella* genus and sulfate reducing bacteria (SRB) were assessed by quantitative polymerase chain reaction (qPCR). While no significant difference in *A. parvulum* abundance was detected between CD, UC and control subjects, CD patients with severe inflammation showed significantly higher abundance of *A. parvulum* as compared to controls and CD subjects with mild inflammation ($p< 0.05$; Figure 3.5.10 A-B). In contrast, no significant difference in *A. parvulum* abundance was detected amongst UC cases of differing disease severities and non-IBD children (Figure 3.5.10 C). The *Veillonella* genus, on the other hand, showed a significant increase in children with UC relative to controls ($p<0.05$; Figure 3.5.10 A). Additionally, the relative abundance of *Veillonella* in children with severe UC was significantly higher as compared to the control group ($p< 0.05$; Figure 3.5.10 C). Finally, no significant difference was detected in the relative abundance of *F. nucleatum* or SRB amongst CD, UC and control subjects (Figure 3.5.10 A).

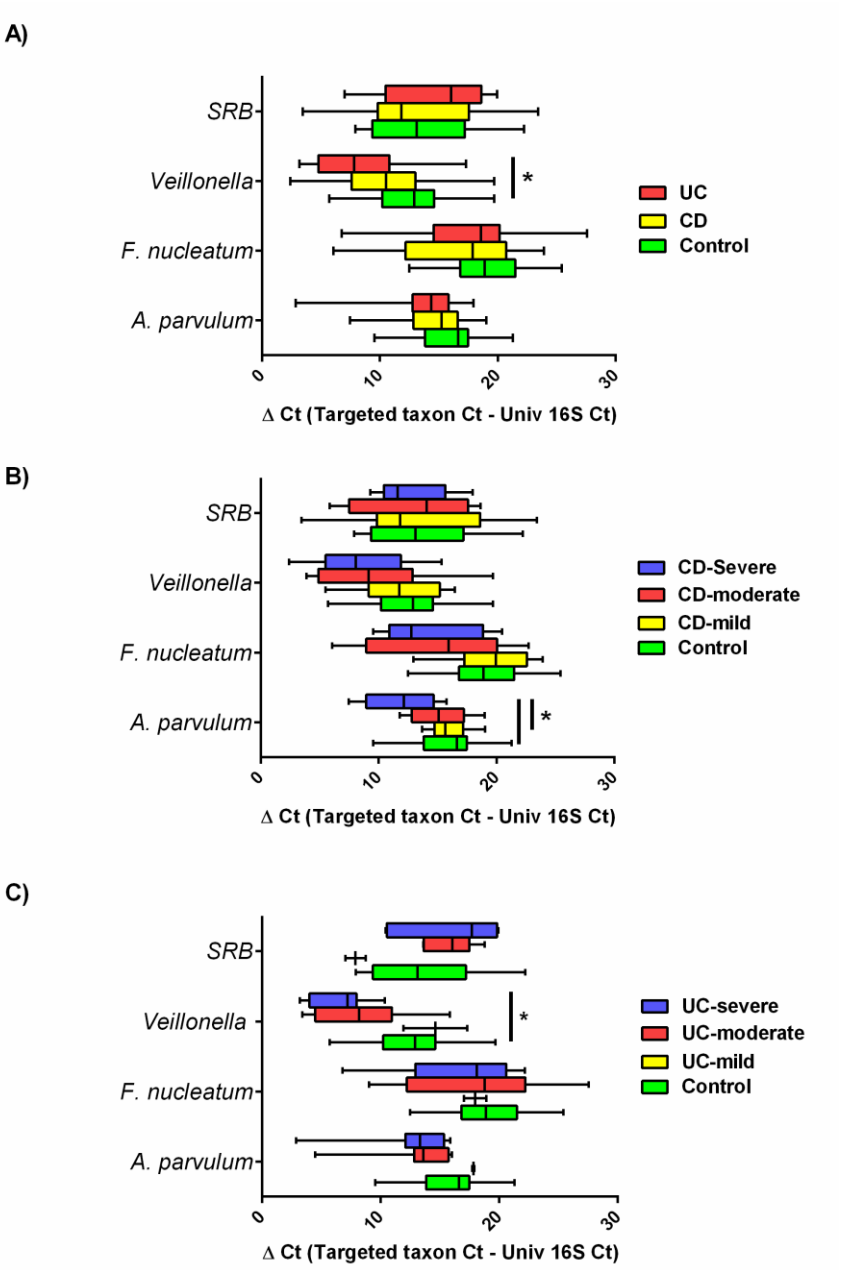


Figure 3.5.10: *Atopobium parvulum* and *Veillonella* are biomarkers of CD and UC severity, respectively, as revealed by qPCR. A, *A. parvulum*, *F. nucleatum*, *Veillonella* genus and sulfate reducing bacteria (SRB) abundance at the RC microbial communities expressed as ΔCt (the numbers of control, CD and UC samples used were 19, 33 and 16 for *A. parvulum*,14, 31 and 14 for *F. nucleatum*, 19, 33 and 16 for *Veillonella* and 17, 32 and 13 for SRB, respectively). ΔCt was calculated by subtracting average Ct values of universal 16S rRNA from average Ct values of the targeted taxon. **B** and **C**, the samples of CD and UC in panel A were categorized as a function of disease severity. The bars represent the median. A Kruskal-Wallis test followed by Dunn's test were used for statistical comparisons (*p < 0.05).

3.6. Characterization of the gut microbiota at the terminal ileum (TI) of pediatric IBD.

3.6.1. Diversity of the terminal ileal (TI) microbiota in pediatric IBD.

The diversity of the TI microbiota was characterized in mucosal aspirates from 53 subjects (15 Controls, 27 CD and 11 UC) by sequencing the V6 hypervariable region of the 16S rRNA gene. A total of 57,435,646 high quality reads were generated with an average of 1,083,691±591,766 reads per sample. Reads were clustered into 4384 OTUs based on 97% identity. The number of OTUs was reduced to 3531 OTUs after filtering out singletons with a Good's coverage of > 99.9% for each sample. The richness and diversity of the terminal ileum microbiota were characterized by estimating the Chao1 and Shannon indices using 499,620 randomly sampled reads from each sample. Neither richness nor diversity of the TI microbiota was altered in either CD or UC in comparison to the control group (Figure 3.6.1 A-B). Moreover, no differential diversity was detected among control and CD patients with either ileal, colonic or ileocolonic inflammation (Figure 3.6.1 C-D).

3.6.2. Dysbiosis of the TI microbiota in pediatric IBD.

To characterize the microbial composition at the TI of CD, UC and control subjects, 53 samples from the three groups were employed (n=27, 11 and 15 for CD, UC and controls, respectively). Analysis of the Illumina HiSeq2500 reads identified 22 phyla in total from all samples (20 phyla from CD samples and 21 from either control or UC samples; Figure 3.6.2 A). More than 99% of the 16S rRNA reads were assigned to only 6 phyla; Actinobacteria, Bacteroidetes, Fusobacteria, Firmicutes, Proteobacteria and Tenericutes (Figure 3.6.2 A). Children with CD showed a higher relative abundance of Proteobacteria (18.12±4% in CD vs 11.5±6.2% in control) in association with a lower abundance of Tenericutes (3.02±1.08% in

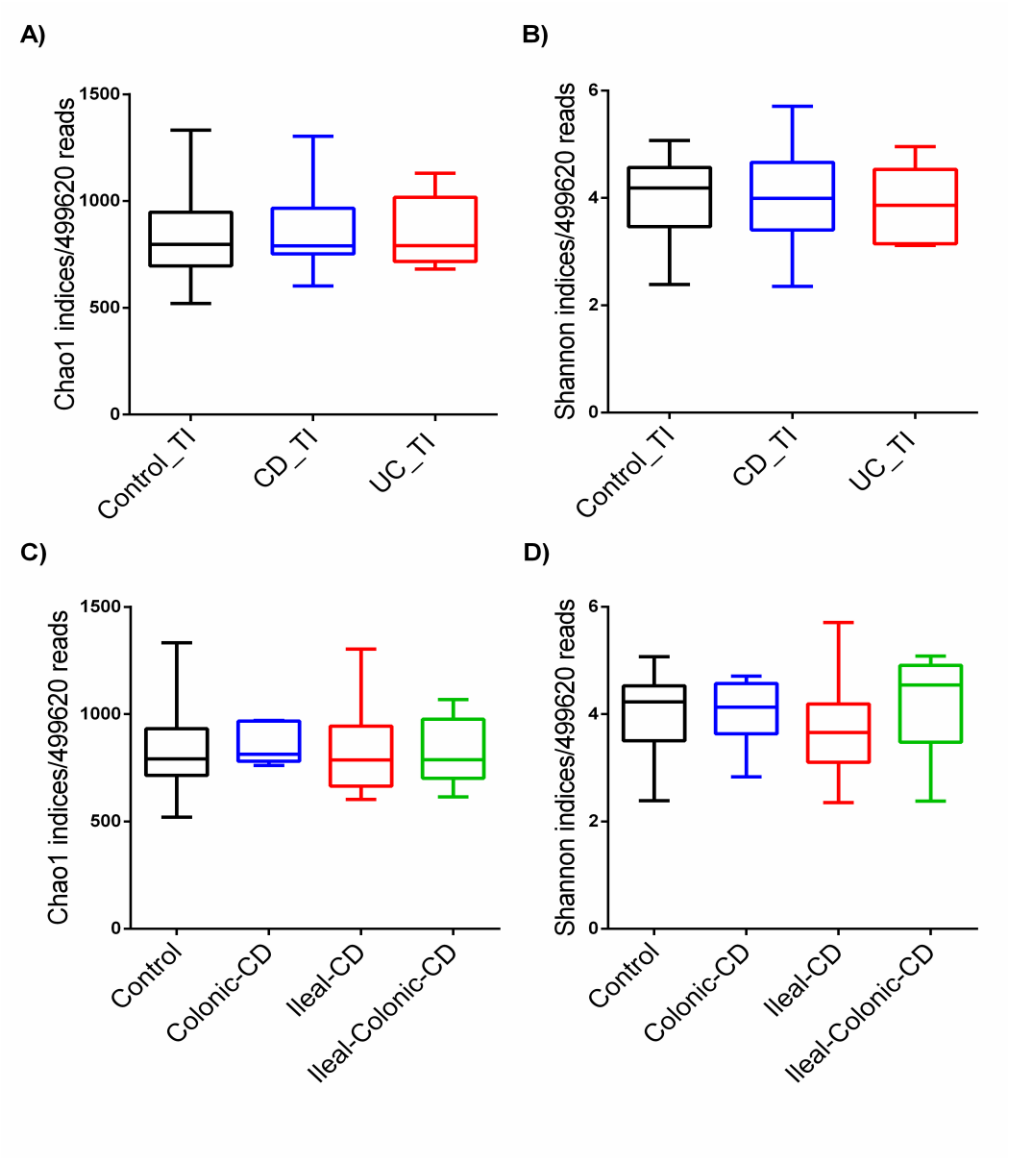


Figure 3.6.1: The gut microbiota exhibit similar diversity at the terminal ileum of pediatric patients with IBD in comparison to control children. Richness and diversity of the gut microbiota were expressed by the Chao1 and Shannon indices as calculated by QIIME 1.7. **A-B**, Diversity indices of TI microbiota in patients with Crohn's disease (CD; n=27), ulcerative colitis (UC; n=10) and control subjects (n= 14). **C-D**, Diversity indices of the TI microbiota in CD samples categorized by the disease location (Colonic CD, n= 6; Ileal CD, n= 12 and Ileal-colonic, n= 9). A Kruskal-Wallis test followed by Dunn's multiple comparisons were applied for statistical comparison.

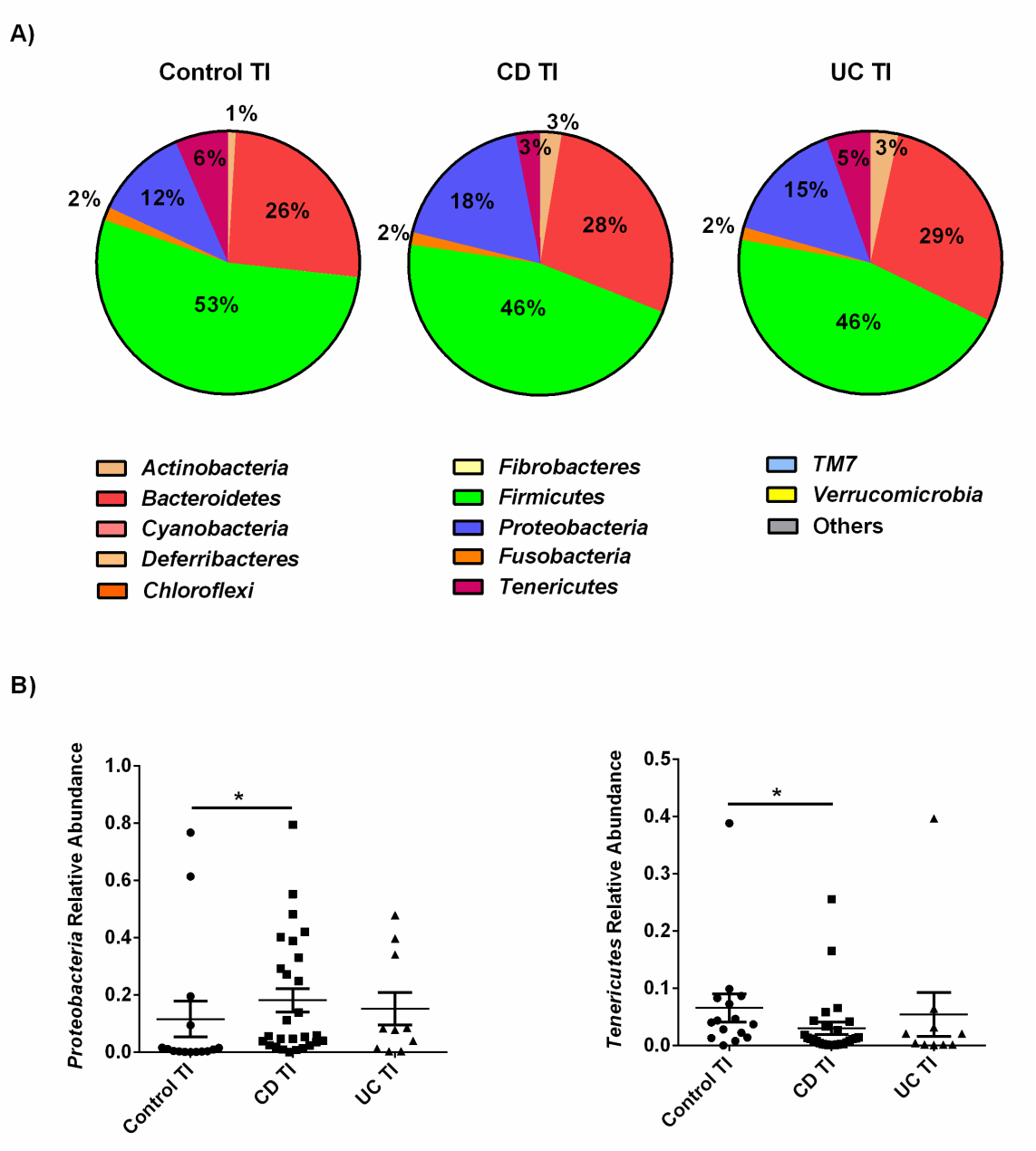


Figure 3.6.2: Taxonomic composition of the terminal ileum microbiota of IBD patients and control subjects. A, Average relative abundance of bacterial phyla identified in patients with Crohn's disease (CD; n=27), ulcerative colitis (UC; n=11) and control subjects (n=15). **B,** Change in relative abundance of *Proteobacteria* and *Tenericutes* in controls, CD and UC patients. A Kruskal-Wallis test followed by Dunn's multiple comparisons were applied for statistical analysis (*p <0.05). The horizontal line represents the mean relative abundance ±SEM in each group.

CD vs 6.5±2.4% in control) as compared to the control microbiota of the TI (Figure 3.6.2 B). At lower phylogenetic levels, 36 taxa exhibited a differential relative abundance as indicated by a non-parametric Kruskal-Wallis test (Bonferroni corrected $p < 0.0167$; Appendix XVI). A linear discriminant effect size analysis (LEfSe) showed that CD was characterized by a depletion of Lachnospiraceae, *Clostridium,* Tenericutes, *Eubacterium rectale, Roseburia, lachnospira, Leptotrichia, Dorea formicigenerans,Prevotella veroralis* and *Blautia producta* (Figure 3.6.3 A). On the other hand, increased abundance of Proteobacteria, Veillonellaceae, Actinobacteria, and *Gemella* were the most discriminant features between CD and control microbial communities (Figure 3.6.3 A). The major effect in the dysbiosis of UC microbiota was the decreased relative abundance of Clostridia, Lachnospiraceae, *Roseburia, Eubacterium rectale, Dorea formicigenerans,* Caulobacteraceae and *Selenemonas* and increased relative abundance of Actinobacteria, *Veillonella,* Neisseriaceae, *Raoutella, Catonella, Neisseria bacilliformis*, *Gemella*, *Kingella oralis* and *Prevotella oulorum* (Figure 3.6.3 B). The relative abundance of differential taxa were further analyzed by PLS-DA modeling which resulted in a detectable and significant separation of control, CD and UC terminal ileum communities (Figure 3.6.3 C). Altogether these results indicate a substantial alteration of the TI microbial community structure in IBD patients as compared to control individuals.

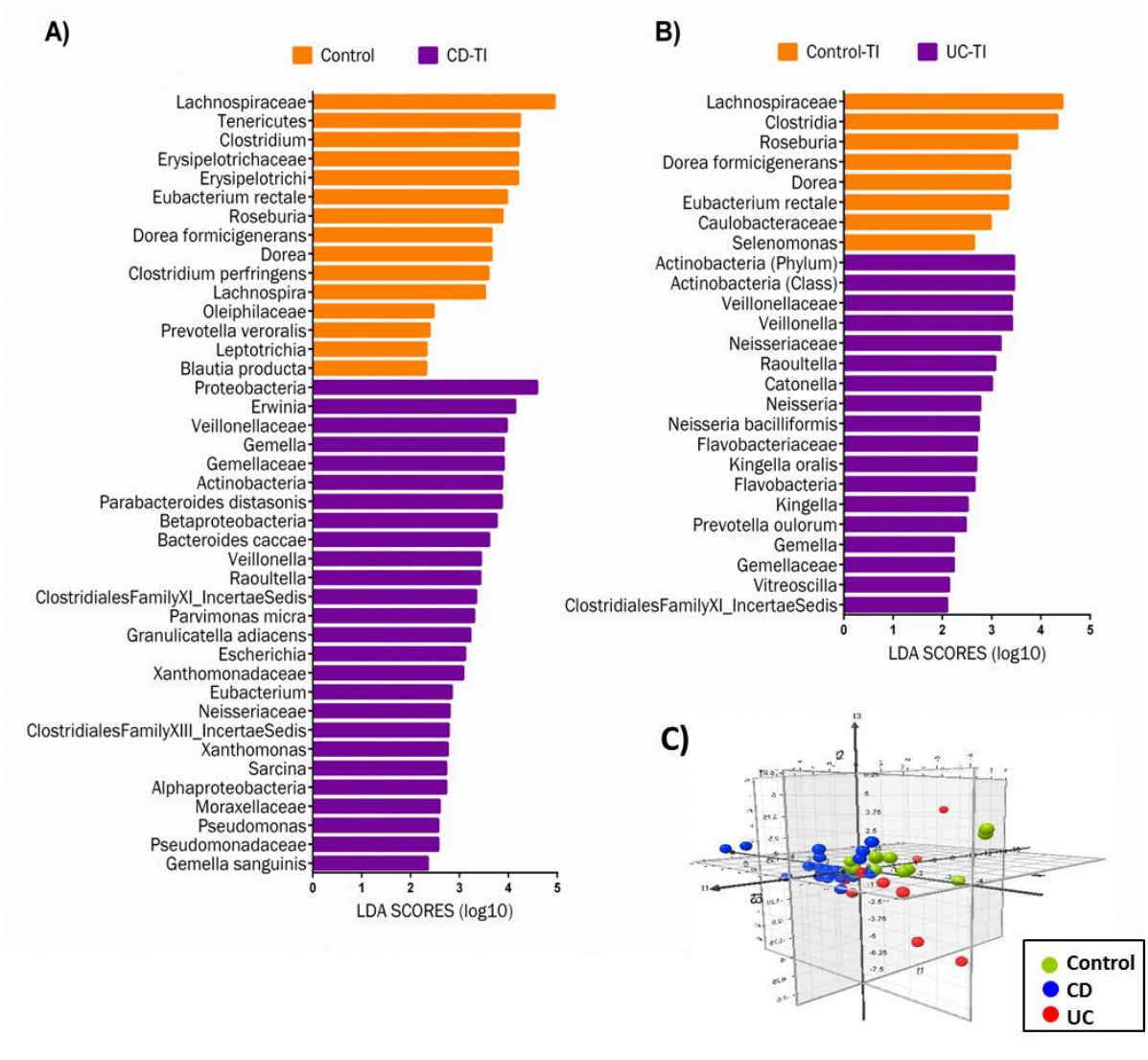


Figure 3.6.3: Dysbiosis of the terminal ileum microbiota in pediatric IBD. The relative abundance of the bacterial taxa obtained from the analysis of the Illumina HiSeq reads were analysed by linear discriminant analysis (LDA) followed by effect size assessment using LEfSe [239]. **A,** Histogram of the LDA effect size score for CD-specific differentially abundant taxa (n=27 and 15 for CD and controls, respectively). **B,** Histogram of the LDA effect size score for UC-specific differentially abundant taxa (n=15 for control and n=11 for UC). **C,** Partial least squares analysis (PLS-DA) of the terminal ileum microbiota from control subjects (n=15) against CD (n=27) and UC patients (n=11). The model was constructed based on the differentially abundant OTUs as determined by Kruskal-Wallis test followed by Bonferroni correction for multiple comparisons.

3.6.3. Depletion of the TI core microbiota in pediatric IBD.

As previously mentioned in sections 3.4.3 and 3.5.3 , the common core microbiota are anticipated to include the health-relevant symbionts [266]. Therefore, we characterized the TI core microbiota that is shared amongst 75% of the individuals in each category. A total of 316 OTUs were shared by at least 75% of samples from the three groups (Control, CD and UC), that represent 91.1% of the total microbial population. CD and UC microbial communities were characterized by a smaller core microbiota as compared to controls with 231, 233 and 270 core OTUs for CD, UC and control subjects, respectively (Figures 3.6.4 and 3.6.5 A). A Kruskal-Wallis test identified 25 OTUs exhibiting differential abundance between at least two of the three groups (CD, UC and controls; Appendix XVII). As compared to control core microbiota, 3 OTUs related to *Bacteroides*, *Parabacteroides distasonis* and unclassified *Ruminococcaceae* were more abundant in CD. In contrast, thirteen OTUs assigned to Lachnospiraceae members (10 OTUs), *Bacteroides*, *Clostridium*, and *Fusobacterium* were depleted in the core microbiota of CD patients as compared to controls (Appendix XVII and Figure 3.6.5 B). Children with UC, on the other hand, showed an increased abundance of 2 core OTUs asigned to the genus *Veillonella* and decreased relative abundance of 5 OTUs assigned to Lachnospiraceae including *Dorea formicigenerans* as compared to the control core microbiota (Appendix XVII and Figure 3.6.5 B). The pairwise comparison between the core microbiota of CD and UC indicated that 2 OTUs exhibited differential abundance. The first OTU, *Vitreoscilla,* was highly abundant in UC, while the second OTU, *Clostridium*, was enriched in CD (Appendix XVII and Figure 3.6.5 B). Furthermore, estimation of the Chao1 and shannon indices revealed a significantly lower richness but not diversity of the 0.75 core microbiota from CD as compared to the control group ($p < 0.001$; Figure 3.6.5 C-D).

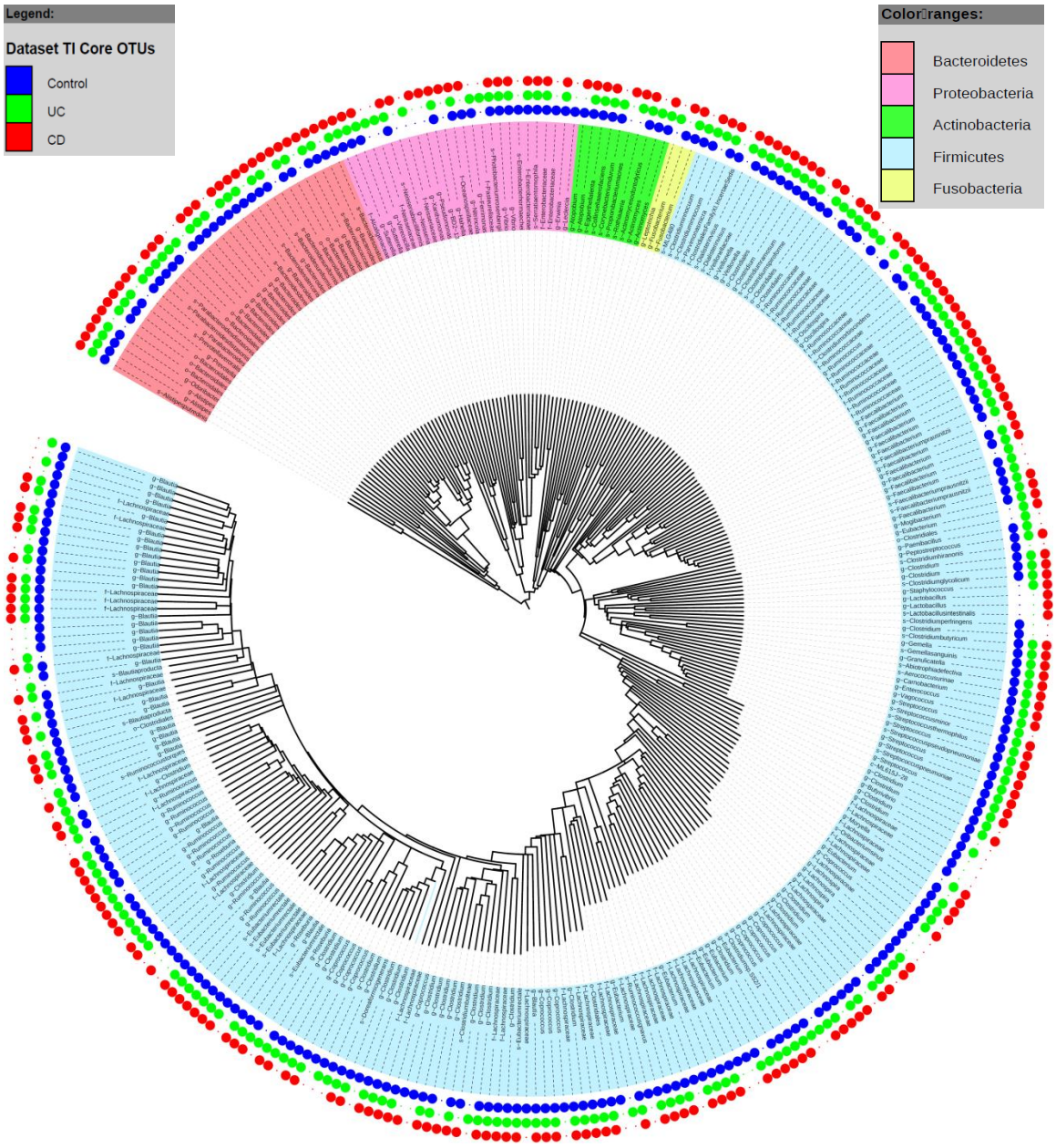


Figure 3.6.4: Phylogenetic tree of the microbial taxa detected in the terminal ileum of at least 75% of the examined samples within each patient cohort. The figure was generated using the iTOL (Interactive Tree of Life) web package (http://itol.embl.de/upload.cgi; [243]). Different phyla are denoted with different colors. Taxa marked with a blue circle were identified as members of the core microbiota of the control subjects. Taxa marked with green or red circles were identified as members of the core microbiota of the UC and CD patients, respectively.

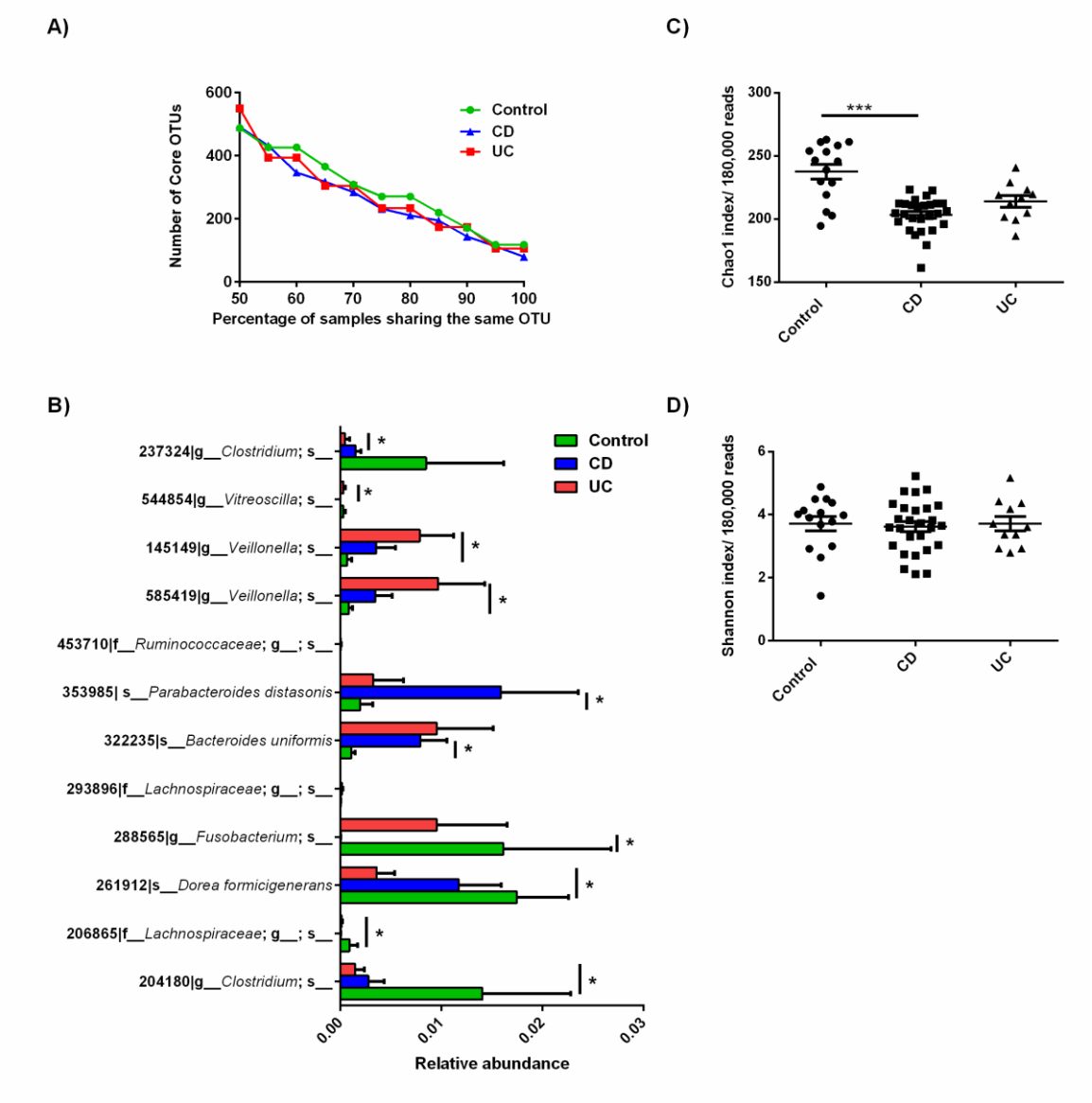


Figure 3.6.5: CD and UC are characterized by a depleted terminal ileum core microbiota as compared to controls. A, The size of the core microbiota shared by different sample percentages of control subjects (green), and CD (blue) and UC (red) patients. OTU corresponds to operating taxonomic unit. **B,** Relative abundance of the statistically significant differential core OTUs among CD, UC and control groups as revealed by Kruskal-Wallis followed by Dunn's post hoc test. **C** and **D**, Chao1 and Shannon indices of 0.75 TI core microbiota, respectively. Sequences assigned to the OTUs that are shared between 75% samples from each category (Control (n=15); CD (n=27) and UC (n=11)) were extracted and the diversity indices were calculated per 180,000 reads from each sample. A Kruskal-Wallis test followed by Dunn's multiple comparisons were applied for statistical analysis (*** p<0.001). The horizontal black lines represent mean±SEM.

3.6.4. The terminal ileal microbiota is associated with pediatric IBD activity.

In order to identify the taxa that may contribute to inflammation, the correlation between the relative abundance of different taxa and the severity of inflammation in either CD or UC patients was examined. A Kruskal-Wallis test identified 37 taxa that exhibited a significant differential abundance between at least two of the three tested groups of CD activity (Mild vs Moderate vs Severe CD; Appendix XVIII; $p < 0.0167$). For UC, a Kruskal-Wallis test revealed only 4 differential taxa correlating with disease activities (Appendix XIX; $p < 0.0167$). PLS-DA model, based on the significantly abundant taxa of either CD or UC, generated 3 clear clusters of the samples dependent on the severity of inflammation (Figure 3.6.6 A-B). A taxa-disease activity biplot correlated *Haemophilus parainfluenza*, *Bacteroides uniformis*, *Clostridium difficile*, *Olsenella* and unclassified *Bifidobacterium* with CD patients with both moderate and severe inflammation. In contrast, *Aeromonas*, *Acidovorax*, *T78*, *Desulfomondales*, *Geobacter*, *Luteibacter rhizovicinus*, *Devosia*, *Acidithiobacillus albertensis*, Hyphomicrobiaceae, *Rhodobacter*, *Selenomonas ruminantium*, *Laribacter hongkongenesis*, *Sphingomonas azotifigens*, *Blautia producta*, *Actinomyces hyovaginalis* and *Mobilincus curtisii* correlated with mild inflammation of CD (Figure 3.6.6 C). *Bifidobacterium adolescentis* and *Turicibacter* correlated with mild UC, while *Capnocytophaga* and its species *C. ochracea* associated with severe UC (Figure 3.6.6 D). In general, these results reveal the association between dysbiosis of the TI microbiota and the severity of inflammation in pediatric IBD.

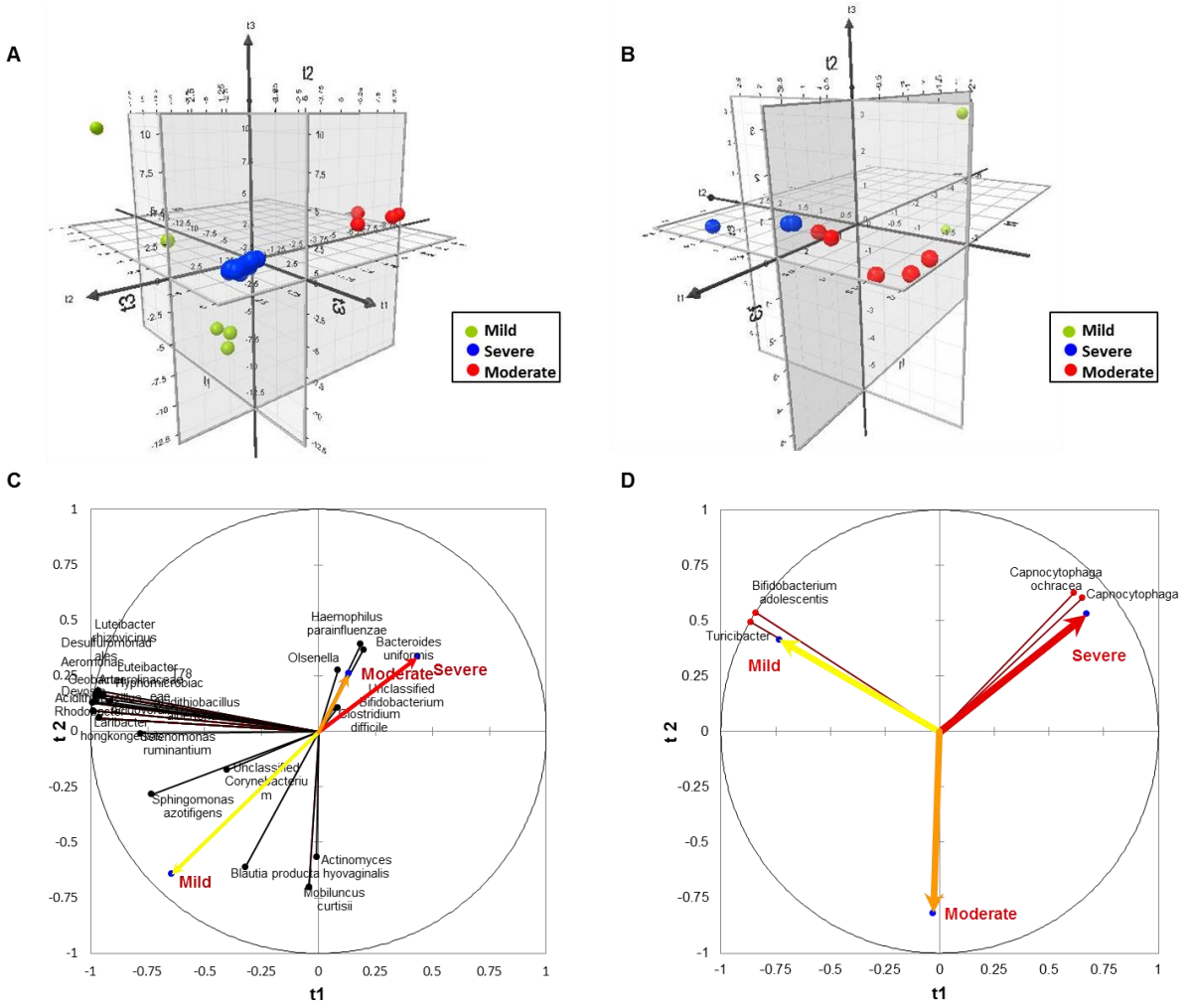


Figure 3.6.6: Partial least squares analysis of the terminal ileum microbiota revealed biomarkers of disease severity. A, PLS-DA of CD patients with severe inflammation (n=18) against CD patients with mild and moderate inflammation (n=5 and 4, respectively). **B,** PLS-DA of UC patients with severe inflammation (n=4) against CD patients with mild inflammation (n=2) and CD patients with moderate inflammation (n=4). The model was constructed based on the differentially abundant taxa as determined by a Kruskal-Wallis test followed by Bonferroni correction for multiple comparisons. The PLS-DA model showed a predictive ability parameter [Q^2 cum] = 0.6 and 0.567, goodness-of-fit parameter [R^2Y cum] = 0.971 and 0.765 and prediction accuracy of 100% and 90.91% for CD and UC, respectively. **C-D,** Biplot analysis of the first two components of the PLS-DA models in panel A and B showing the significant taxa relative to disease activity (arrows).

3.7. Alteration of the identified CD-specific microbial biomarkers across age, gender and intestinal site.

Prior studies have revealed age and gender-related changes in the composition of gut microbiota [100, 102, 267]. Therefore, we aimed to investigate the pattern of the identified CD-altered taxa across age and gender. This analysis was not performed with the UC cohort due to an insufficient number of samples. First, we compared the relative abundance of each taxon in children of 12 years old and under to those above 12 years old. Enterobacteriaceae, Gammaproteobacteria and Proteobacteria were the only taxa that showed significant difference in their relative abundance at the terminal ileum between the 2 age groups as they were only associated with children under 12 years old (Figure 3.7.1 A). Moreover, no significant difference was detected between the relative abundance of CD biomarkers as a function of gender (Figure 3.7.1 B).

As previously mentioned in section 1.4, the structure of the gut microbiota varies along the intestinal tract. So, we intended to characterize the gut microbiota structures at different regions of the intestinal tract in CD individuals. We selected only 20 CD subjects for this analysis and sampled from three different intestinal locations (LC, RC and TI). A Kruskal–Wallis test followed by a Dunn's test for multiple comparisons identified 25 taxa exhibiting differential abundance between the three tested sites (Figure 3.7.1 C). The identified 25 taxa, namely Actinomycetaceae, *Actinomyces*, *Actinomyces odontolyticus*, *Actinomyces oris*, Propionibacteriaceae, *Propionibacterium*, Aerococcaceae, *Aerococcus sanguinicola*, Carnobacteriaceae, *Granulicatella*, *Granulicatella adiacens*, Paenibacillaceae, *Paenibacillus*, Streptococcaceae, *Streptococcus*, *S. infantis*, *S. pseudoneumoniae*, *S. thermophilus*, Bradyrhizobiaceae, *Rhodopseudomonas*, *Leptothrix*, *Leptothrix discophora*, *Lautropia*, Clostridiaceae and *Actinobacillus pleuropneumoniae,* were significantly enriched at the TI as compared to the RC and/or LC of CD patients (Figure 3.7.1 C).

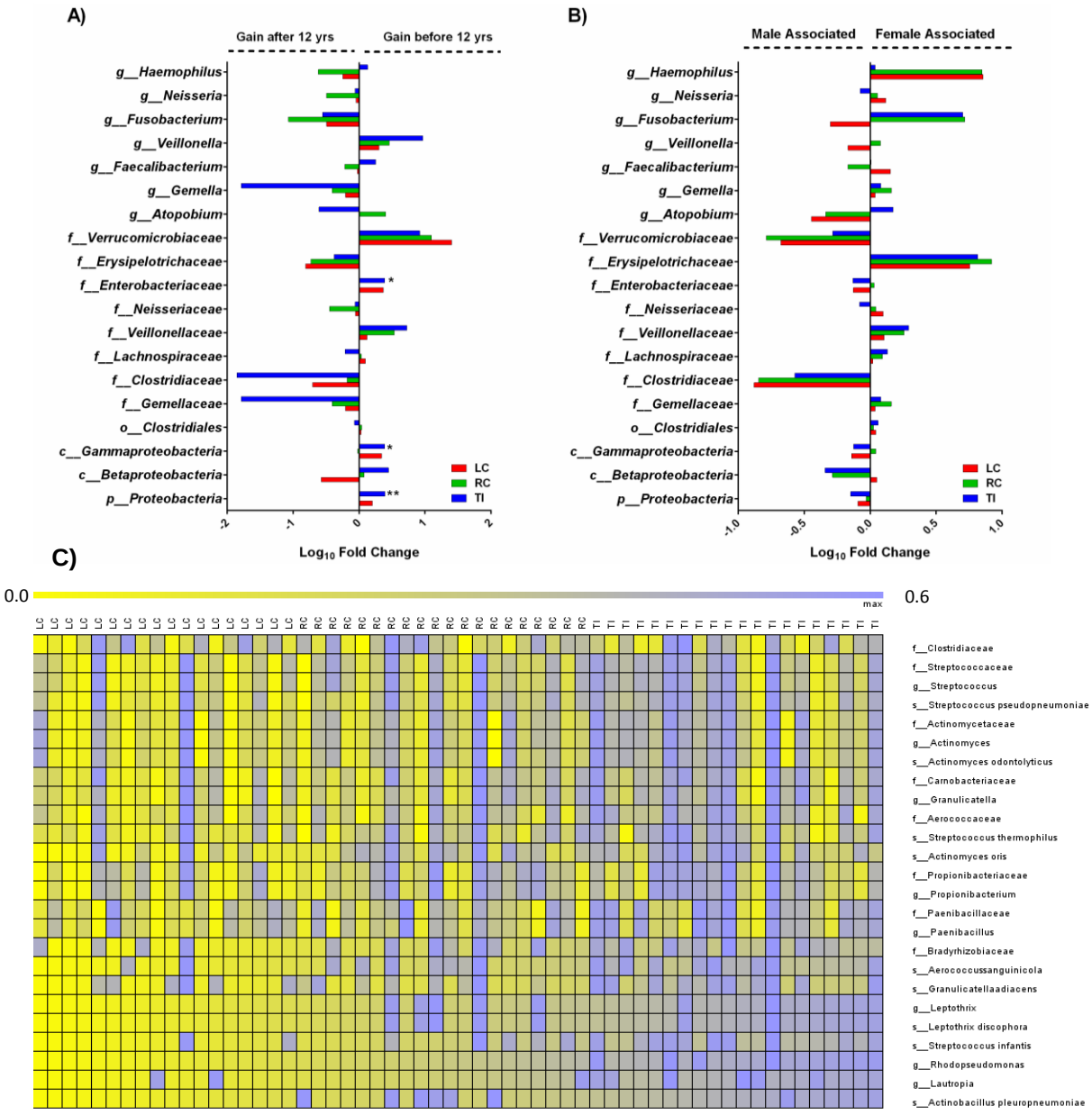


Figure 3.7.1: Alteration of the CD specific microbiota as a function of age, gender or intestinal location. A, Log_{10} fold change of CD-biomarkers in children ≤ 12 years against children > 12 years old at the three tested intestinal sites. The fold change was calculated by dividing the average relative abundance of each taxon in subjects less than 12 years old by the average relative abundance in children more than 12 years old. **B,** Log_{10} fold change of CD-biomarkers as a function of gender. The fold change was calculated by dividing the average relative abundance of each taxon in females by those in males. Two way ANOVA followed by Sidak's test were used for statistical comparisons (*p< 0.05 and **p< 0.01). **C,** Heat map illustrating the relative abundance of differentially abundant taxa between left colon (LC), right colon (RC) and terminal ileum (TI) of CD individuals (n=20).

3.8. Functional dysbiosis of gut microbiota along the intestine of new onset pediatric IBD.

While metagenomic shotgun sequencing is the best approach to determine the functionality of microbial communities, the phylogenetic information strongly correlates to biomolecular function [66]. Therefore, we predicted the metabolic functionality of the gut microbial communities from the 16S rRNA phylogenetic information characterized from IBD and healthy children using the PICRUST approach [244]. PICRUST is a technique that infers the metagenomic data from the 16S rRNA information and a reference genome database through evolutionary modeling [244]. A Kruskal-Wallis test of the generated relative abundance of the KEGG pathways identified 16, 51 and 12 pathways with significant alterations in their relative abundance among CD, UC and controls at the LC, RC and TI, respectively (Appendices XX-XXII and figures 3.8.1 and 3.8.2). The microbial community at the LC of CD patients was characterized by increased relative abundance of bacterial secretion systems and the metabolism of D-Alanine, pyruvate and retinol and decreased relative abundance of carbohydrate metabolism, glycolysis/gluconeogenesis and pantothenate and CoA synthesis as compared to control LC communities (Appendix XX and figure 3.8.1A). No significant difference was detected between the microbial functionalities in UC and control children at the LC. At the RC, bacterial secretion system and propanoate metabolism increased, while pentose phosphate pathway, primary bile acid biosynthesis and xylene degradation decreased in CD as compared to non-IBD controls (Appendix XXI and figure 3.8.2). The microbial communities at the RC of UC subjects showed increased relative abundance of aminobenzoate degradation, arachidonic acid metabolism, bacterial motility proteins, bacterial secretion system, biosynthesis and biodegradation of secondary metabolites, carbohydrate digestion and absorption, cell motility and secretion, chloroalkane

and chloroalkene degradation, citrate cycle (TCA cycle), D-Arginine and D-ornithine metabolism, ethyl benzene degradation, glutathione metabolism, indole alkaloid biosynthesis, inorganic ion transport and metabolism, lipopolysaccharide biosynthesis proteins, lipopolysaccharide biosynthesis, metabolism of cofactors and vitamins, nitrogen metabolism, porphyrin and chlorophyll metabolism, propanoate metabolism, protein folding and associated processing, RNA transport, *Staphylococcus aureus* infection, toluene degradation, ubiquinone and other terpenoid-quinone biosynthesis, ubiquitin system and valine, leucine and isoleucine degradation (Appendix XXI and figure 3.8.2). On the other hand, the relative abundance of glycerolipid metabolism, lysine biosynthesis, methane metabolism, nucleotide excision repair, pentose phosphate pathway, primary bile acid biosynthesis, secondary bile acid biosynthesis, sporulation, starch and sucrose metabolism, thiamine metabolism were decreased in the RC microbial community of UC relative to the controls (Appendix XXI and figure 3.8.2). The microbial communities at the TI of CD patients were characterized by increased chaperons and folding catalyst, flavone and flavonol biosynthesis, lipid biosynthesis proteins, taurine and hypotaurine metabolism and toluene degradation and depleted bacterial chemotaxis, fructose and mannose metabolism and pentose phosphate pathway as compared to the control TI microbiome (Appendix XXII and figure 3.8.1B). Finally, no significant difference in the microbial functionalities was detected between the TI of UC and the control group.

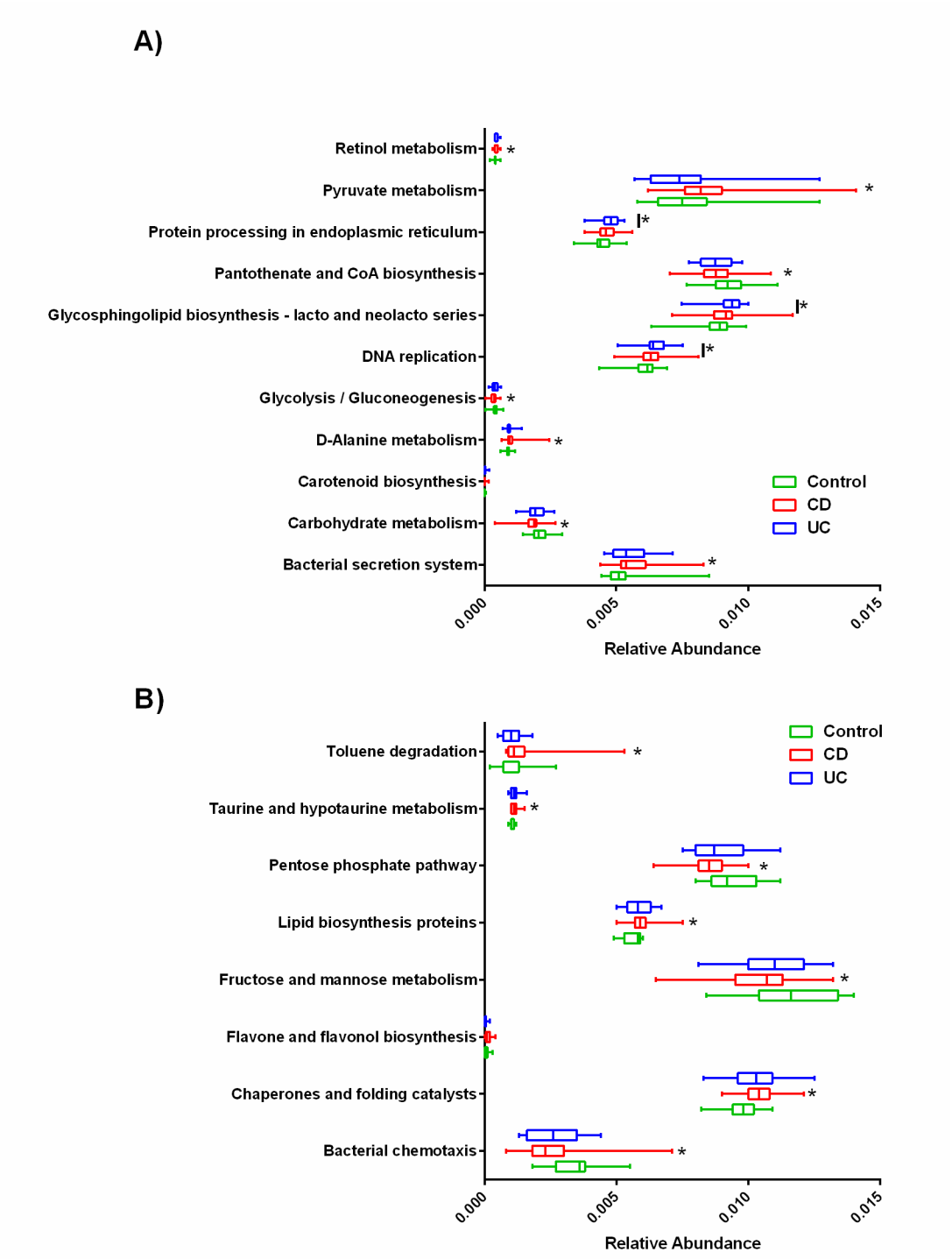


Figure 3.8.1: Microbial metabolic pathways with significant differential abundance in the LC (A) and TI (B) communities of new onset pediatric IBD patients. Relative abundance of KEGG pathways were predicted by PICRUST approach [244]. Statistical comparisons were performed using Kruskal-Wallis test followed by Bonferroni correction for multiple comparisons. Bonferroni corrected significance level is 0.0167.

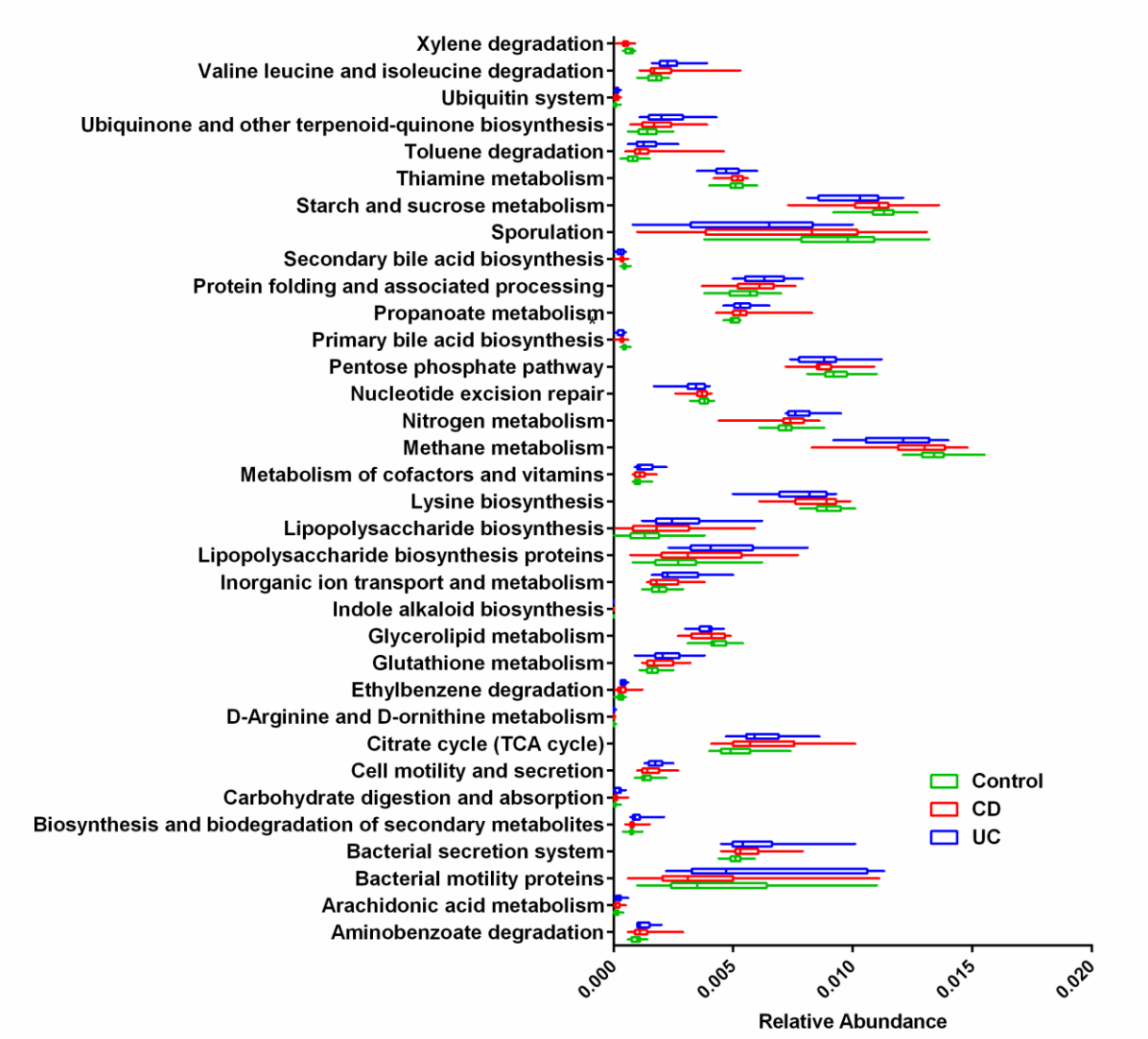


Figure 3.8.2: Microbial metabolic pathways with significant differential abundance in the RC communities of new onset pediatric IBD patients. Relative abundance of KEGG pathways were predicted by PICRUST approach [244]. Statistical comparisons were performed using Kruskal-Wallis test followed by Bonferroni correction for multiple comparisons. Bonferroni corrected significance level is 0.0167.

3.9. *Atopobium parvulum* modulates the severity of pediatric IBD inflammation in presence of other gut microbes.

3.9.1. *Atopobium parvulum* induces rapid and severe pan-colitis in 129/SvEv *Il-10*$^{-/-}$ mice under specific pathogen free (SPF) conditions.

The results of the gut microbiota characterization showed that *A. parvulum* is a biomarker of CD severity. To functionally test the colitogenic potential of *A. parvulum*, we utilized the colitis susceptible 129/SvEv *Il-10*$^{-/-}$ knockout mice [253, 268]. Germ-free *Il-10*$^{-/-}$ mice were transferred to specific pathogen free (SPF) housing and gavaged with *A. parvulum* (10^8 CFU/mouse/week) for 6 weeks. *A. parvulum*-infected *Il-10*$^{-/-}$ mice exhibited a shorter colon as compared to unifected mice at the time of necropsy (Figure 3.9.1 A). Additionally, mice colonoscopy illustrated ulcerated colonic walls as evidence of macroscopic inflammation in *A. parvulum*-infected *Il-10*$^{-/-}$ mice, while control uninfected *Il-10*$^{-/-}$ mice maintained healthy colonic mucosa (Figure 3.9.1 B). Representative histological sections from the cecum and distal colon of *A. parvulum*-associated *Il-10*$^{-/-}$ mice showed crypt abscesses and hyperplasia, submucosa expansion with inflammatory infiltrate, goblet cell depletion and submucosal edema. Uninfected control *Il-10*$^{-/-}$ mice, on the other hand, displayed minimal signs of inflammation (Figure 3.9.1 C). Therefore, histological inflammation scores were significantly higher in *A. parvulum* associated *Il-10*$^{-/-}$ mice as compared to uninfected mice ($p<0.05$; Figure 3.9.1 D). Furthermore, qRT-PCR indicated a significantly higher expression of *Il-12* and *Il-1β* and a high but non-significant level of *Cxcl-1* and *Il-17* in *A. parvulum* colonized *Il-10*$^{-/-}$ mice as compared to uninfected mice ($p<0.05$; Figure 3.9.2 A-D). Overall, these results indicate that *A. parvulum* induces severe pancolitis from the cecum to the colon in a genetically susceptible murine model of IBD.

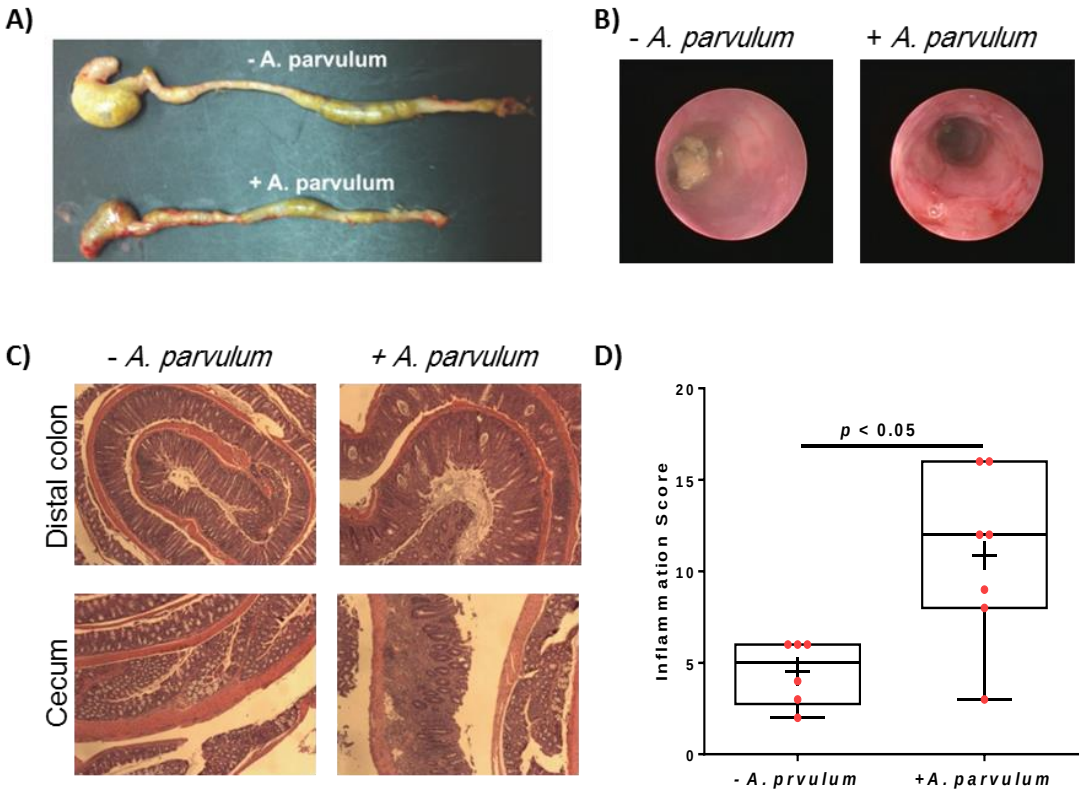


Figure 3.9.1: *Atopobium parvulum* induces rapid and severe pan-colitis in *Il-10*$^{-/-}$ mice. **A,** The cecum and colon from specific pathogen free (SPF) *Il-10*$^{-/-}$ mice that were either associated or not-associated with *A. parvulum* were visually observed. **B,** Colon inflammation was monitored macroscopically with a murine endoscope (data is representative of 2 animals). **C,** Representative histological sections of the distal colon and cecum. **D,** Blinded histological score of inflammation (n=6 to 7 per group; horizontal lines indicate mean and crosses indicate median; comparison by Mann-Whitney two tailed test).

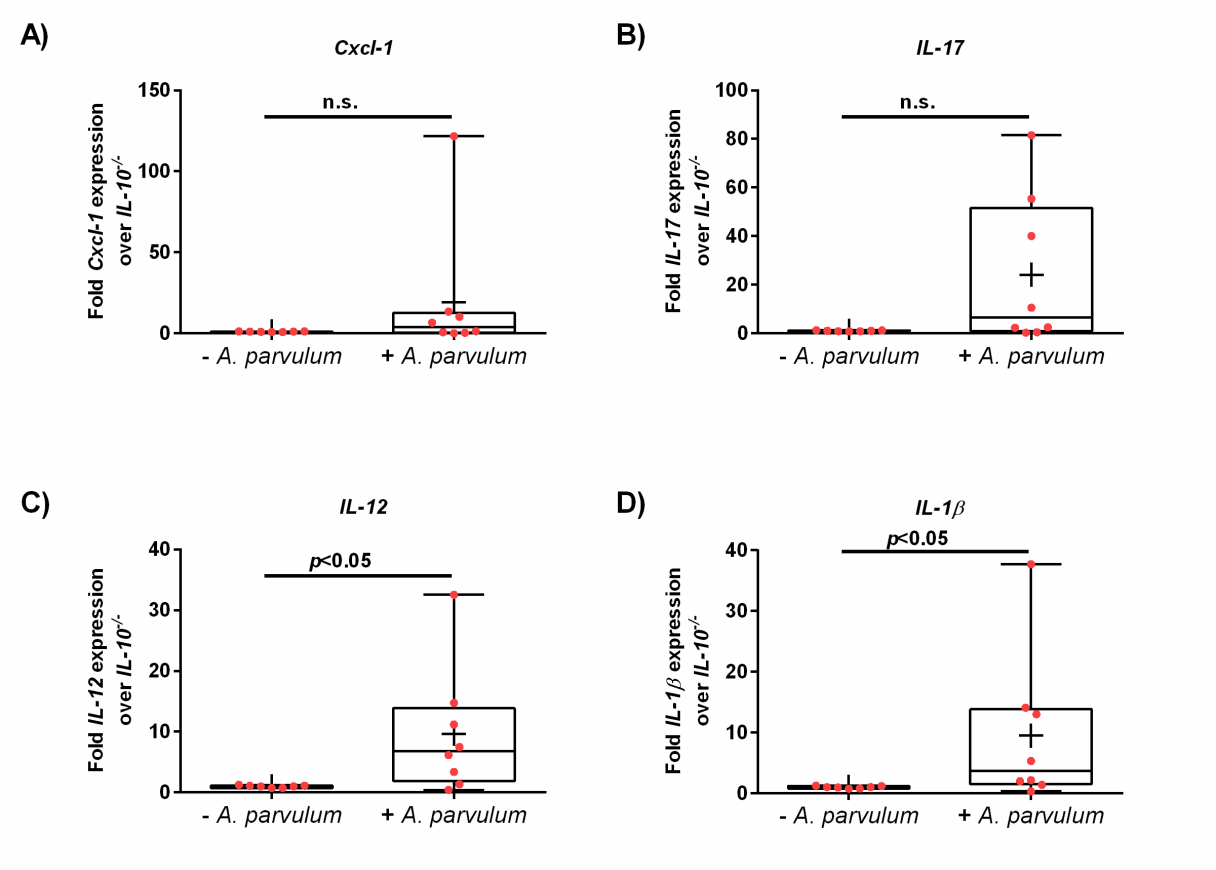


Figure 3.9.2: *A. parvulum* induces cytokine expression in conventionalized *Il-10*$^{-/-}$ mice. 129/SvEv, *Il-10*$^{-/-}$ mice were associated or not with *A. parvulum* and kept under SPF conditions (n=7 to 8 per group). Total RNA was extracted from colonic intestinal tissues 6 weeks post-association and *Cxcl1*, *Il-17*, *Il-12* and *Il-1β* expression was measured by qRT-PCR, panels **A** to **D**. Error bars indicate the mean±SEM. A Mann-Whitney U test was performed to assess statistical significance. n.s.; non-significant.

3.9.2. Bismuth alleviates *A. parvulum*-induced colitis.

To test the causative role of H_2S in colitis, we assessed whether an H_2S scavenger (bismuth) could prevent colitis or reduce inflammation induced by *A. parvulum* in an *Il-10*$^{-/-}$ mouse model. Consistent with the first cohort, *Atopobium*-associated mice kept under SPF conditions showed macroscopically and histologically severe colitis as indicated by colonoscopy visualization and increased inflammation scores (Figure 3.9.3 A-B). *A. parvulum*-associated mice that received bismuth treatment maintained a healthy colonic wall (Figure 3.9.3 A) with a significant decrease in inflammation score (p=0.007; Figure 3.9.3 B). *Atopobium*-colonized mice also had a higher number of gut associated lymphoid tissue (GALT) foci as compared to SPF mice (p=0.012). However, bismuth treatment did not prevent GALT formation (Figure 3.9.3 C). We also assessed the effect of bismuth on the gut microbial composition of *Atopobium*-associated and unassociated SPF *Il-10*$^{-/-}$ mice. PCA analysis of the gut microbiota composition revealed a significant alteration in the microbial profile of the *Atopobium*-associated mice as compared to *Atopobium*-free SPF mice, while bismuth treatment shifted the microbiota of these 2 groups of mice toward more similar microbial profiles (Figure 3.9.3 D). A Kruskal-Wallis analysis of the identified taxa in the four tested categories revealed a significant differential abundance in 100 taxa (Appendix XXIII). For example, *A. parvulum* colonized mice exhibited a lower relative abundance of Fusobacteria, Cyanobacteria, Betaproteobacteria, Eubacteriaceae, *Eubacterium*, *Ruminococcus* and *Faecalibacterium*, in association with enrichment of Bacteroidaceae and Erysipelotrichaceae as compared to *A. parvulum*-unassociated SPF mice (Appendix XXIII; Bonferroni corrected $p < 0.008$). These results indicate that *A. parvulum* colonization leads to major changes in the composition of the gut microbiota with a significant decrease in the

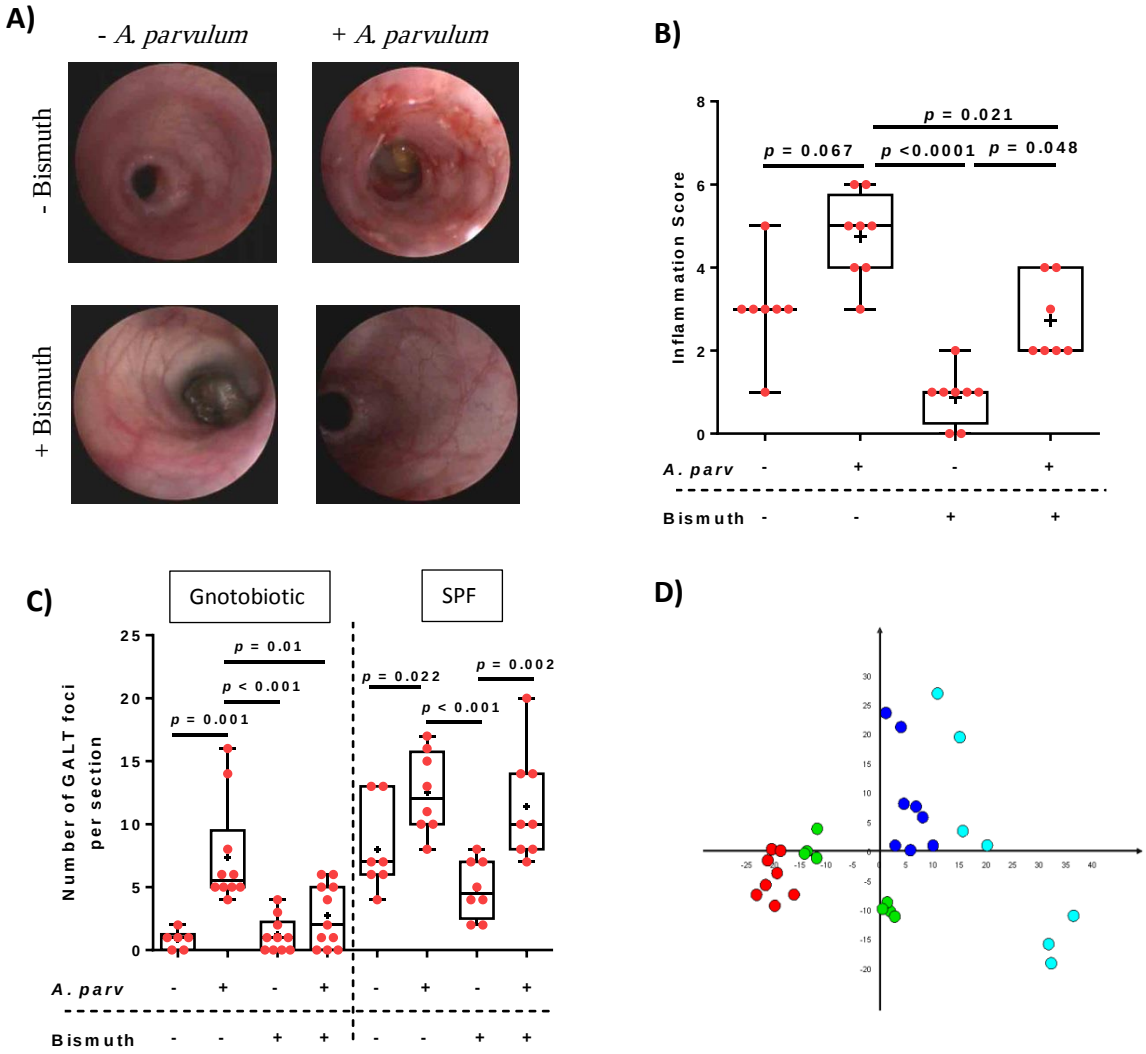


Figure 3.9.3: Bismuth prevents *Atopobium*-induced colitis. A, Representative murine endoscopies of *Il-10*$^{-/-}$ mice associated or not-associated with *A. parvulum*, treated or not-treated with bismuth and kept under SPF conditions for 6 weeks. **B,** Blinded inflammation scores (n=7 to 8 per group) for *Il-10*$^{-/-}$ mice under SPF conditions. **C,** Number of GALT foci of *Il-10*$^{-/-}$ mice associated or not with *A. parvulum* and treated or not with bismuth, and kept under gnotobiotic or SPF conditions (n=6 to 11 per group). **B** and **C** Horizontal lines indicate means and crosses indicate median. Statistical significance was assessed using a Kruskal-Wallis test with a Dunn's post hoc test. **D,** PCA of microbiota from *Il-10*$^{-/-}$ mice kept under SPF conditions (light blue), bismuth-treated (blue), associated with *A. parvulum* (red), and associated with *A. parvulum* and treated with bismuth (green).

abundance of major butyrate-producers. Furthermore, bismuth administration appears to restore a healthier microbiota community.

3.9.3. Gut commensal microbiota are essential for *Atopobium parvulum*-induced colitis development.

To assess the role of commensal microbes in *A. parvulum*-induced colitis development, germ-free mice were mono-associated with *A. parvulum* and were maintained under gnotobiotic conditions. While these mice displayed increased crypt hyperplasia and increases in GALT foci (Figure 3.9.4 A-B), they had no signs of edema, submucosal expansion, ulcers, goblet cell depletion or immune cell infiltration (Figure 3.9.4 A). Bismuth treatment prevented GALT development in mice mono-associated with *A. parvulum* (Figure 3.9.4 B). However, qPCR indicated significantly lower colonization levels of *A. parvulum* in bismuth-fed mice as compared to mice not fed bismuth (Figure 3.9.4 C). To confirm these results, we utilized a chemically induced colitis model of IBD. Swiss Webster germ free mice were gavaged with *A. parvulum* or *Bifidobacterium longum* (10^8 CFU/mouse) and colitis was induced with the addition of 3.5% dextran sodium sulfate (DSS) in their drinking water for 5 days. No differential weight loss was detected after the DSS treatment between *A. parvulum*- and *B. longum*-monoassociated mice (Figure 3.9.5 A). At the time of necropsy, both colon weight and length of *A. parvulum*-colonized mice were significantly lower than *B. longum* associated mice (Figure 3.9.5 B-C). The assessment of 32 cytokines and inflammation markers, secreted by colonic specimens, pointed out a significant increase in the level of only IL-15 in *A. parvulum*-associated mice as compared to *B. longum*-colonized mice (Figure 3.9.5 D). Overall, these results indicate that the gut microbiota are essential for the development of *A. parvulum*-induced colitis.

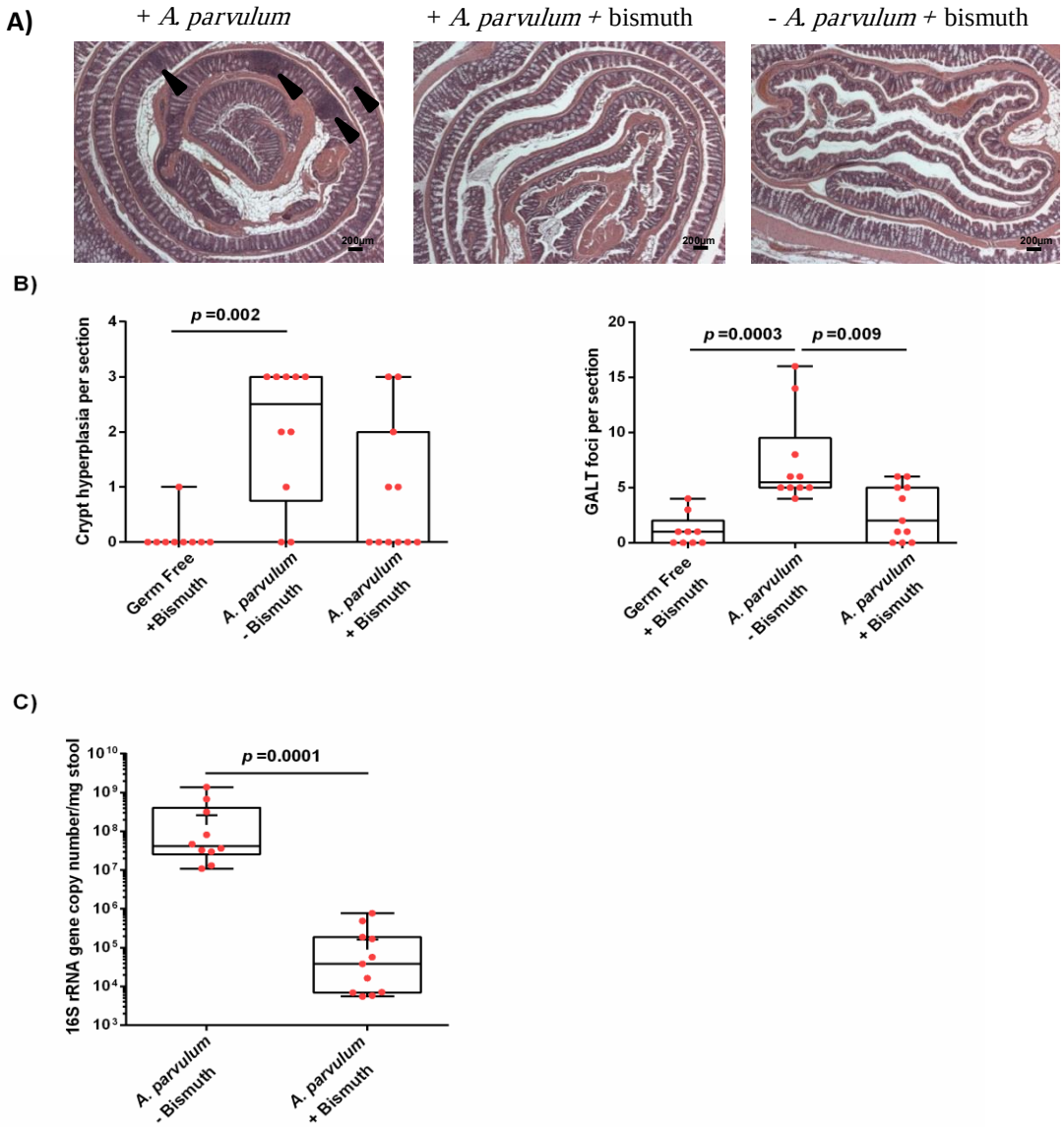


Figure 3.9.4: *Il-10*$^{-/-}$ mice mono-associated with *A. parvulum* exhibit increased hyperplasia scores and GALT foci and bismuth decreases *A. parvulum* colonization level. A, Representative histological sections of distal colon showing gut associated lymphoid tissues in mice monoassociated with *A. parvulum* in the absence of bismuth (GALT; black arrows). **B,** Hyperplasia score and GALT foci of 129/SvEv *Il-10*$^{-/-}$ mice mono-associated or not monoassociated with *A.parvulum* and treated or not treated with bismuth kept under gnotobiotic conditions (n=6 to 11 per group). **C,** Metagenomic DNA was extracted from stool pellets obtained 6 week after mono-association or no mono-association of 129/SvEv *Il-10*$^{-/-}$ mice with *A. parvulum*. Colonization level was estimated using real-time qPCR and reported as the number of 16S rRNA gene copies per mg of stool. For panels **B** and **C,** horizontal lines indicate the mean and error bars indicate SEM. Statistical significance was assessed using a Mann-Whitney U-test.

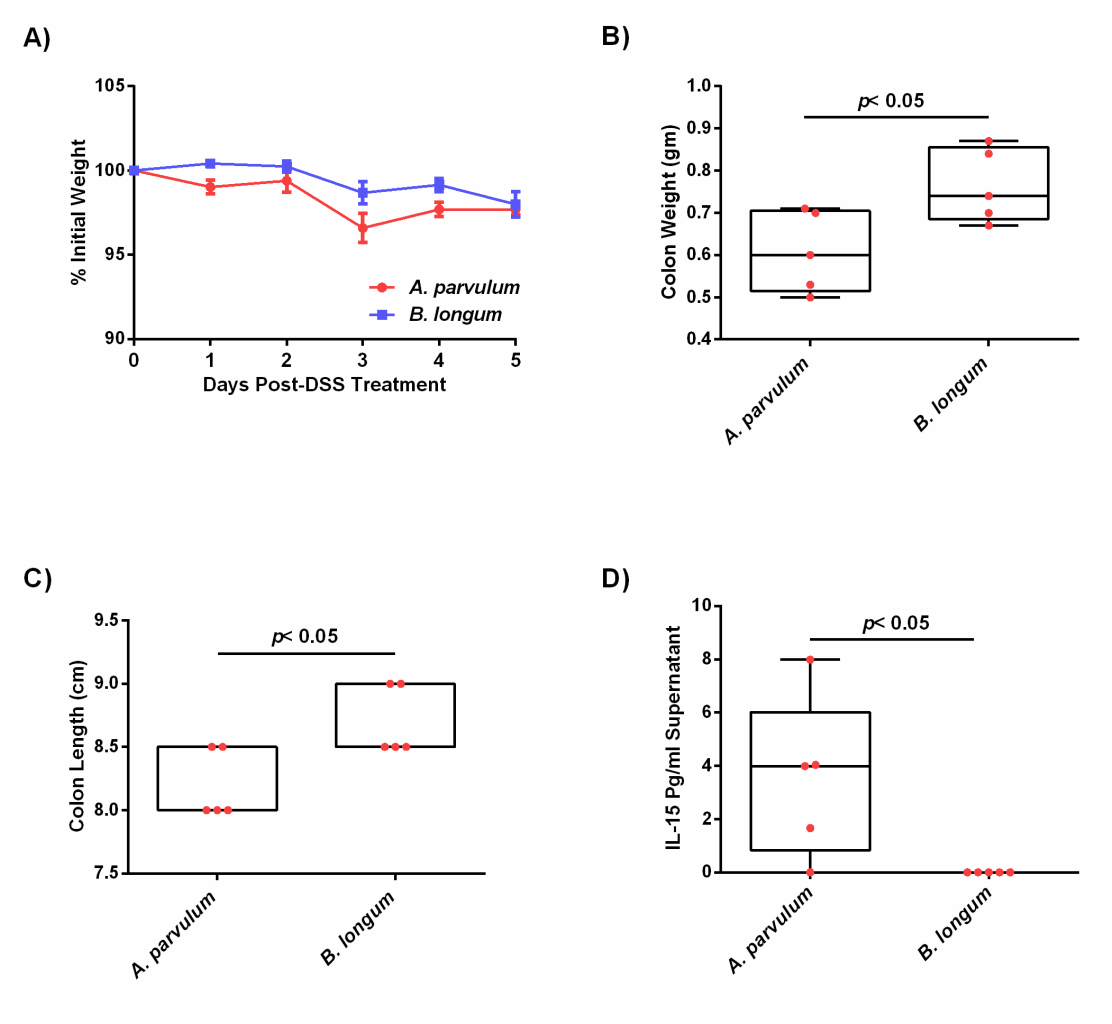
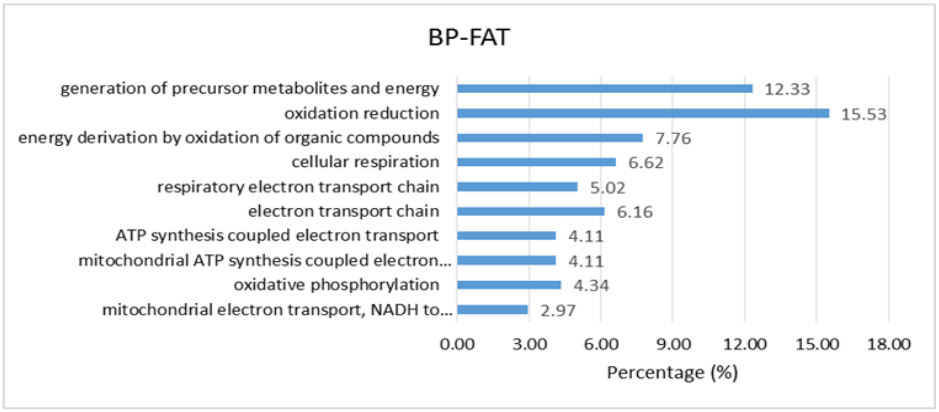


Figure 3.9.5: Gnotobiotic mice mono-associated with *A. parvulum* exhibit an incomplete picture of inflammation post DSS treatment. Swiss Webster germ free mice (4-6 weeks) were monoassociated with *A. parvulum* (n=5) or *B. longum* (n=5). Colitis was induced by adding dextran sodium sulfate (DSS) to the drinking water (3.5%) for 5 days then euthanized with CO_2 asphyxiation. **A,** mice weight during DSS treatment expressed as percentage of initial weight. **B-C**, colon weight and length at the time of necropsy. **D,** Il-15 concentration in the supernatant of 24 hr culture of colonic specimens in RPMI medium as determined by Multiplexing LASER Bead Technology. A Mann-Whitney U test was performed to assess statistical significance. Horizontal lines indicate the mean±SEM.

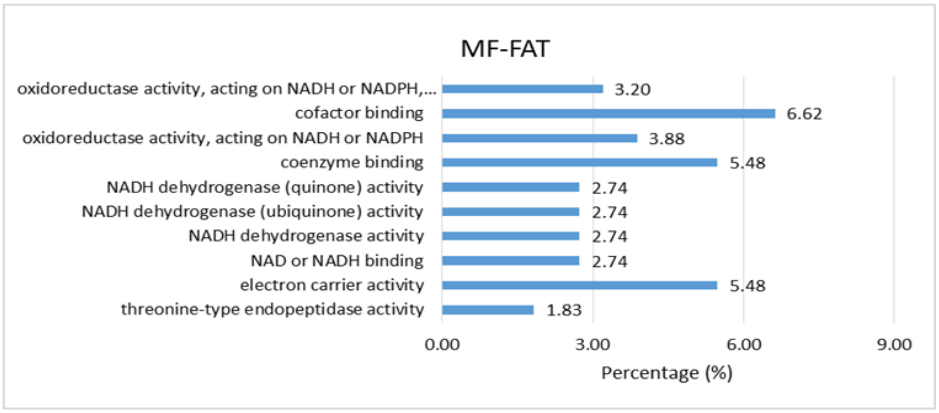
3.10. Alteration of human colonic metaproteome in pediatric IBD.

To obtain mechanistic insights into the role of gut microbes in the severity of IBD, Dr. Daniel Figeys' lab conducted an unbiased, quantitative proteomic analysis of mucosal biopsies from CD subjects of varying disease severities (n=21) and controls (n=8). We identified 3880 proteins of which 490 were differentially expressed among the 3 major groups: severe vs. moderate vs. control (one-way ANOVA with p <0.05; Appendix XXIV). Functional annotation of the differentially expressed genes identified oxidation-reduction, energy generation, oxidative phosphorylation, cellular respiration and electron transport as the most significantly enriched biological processes in CD as compared to control mucosal tissues (Figure 3.10.1 A). Ninety five mitochondrial proteins were identified as major discriminant features, which represent 21.7% of all differentially expressed proteins (Appendix XXIV; Figures 3.10.1 and 3.10.2 A). Proteins potentially driving disease activity were identified by PLS-DA and the analysis of their variable importance projection (VIP) scores (Appendices XXIV-XXV and Figure 3.10.2 B). Notably, components of the mitochondrial hydrogen sulfide detoxification complex [172] were found to be the main proteins driving the separation based on disease severity (Appendix XXV). These proteins, namely sulfur dioxygenase (ETHE1), thiosulfate sulfur transferase (TST), and the components of complexes III and IV of the mitochondrial respiratory chain, were down regulated in CD patients compared to controls (p<0.05). Secondary validation by qRT-PCR confirmed the repression of the TST transcript (5 fold decrease; p<0.05) in CD and UC patients (Figure 3.10.2 C). Moreover, the expression levels of the cytochrome c oxidase subunit IV and the sulfide dehydrogenase genes (SQR), which also contribute to the detoxification of H_2S, were significantly down-regulated in CD and/or UC patients, as measured by qRT-PCR (Fig. 3.10.2 D-E; p< 0.05). These findings indicate that transcriptional regulation contributes to the observed abundance changes of H_2S

A)



B)



C)

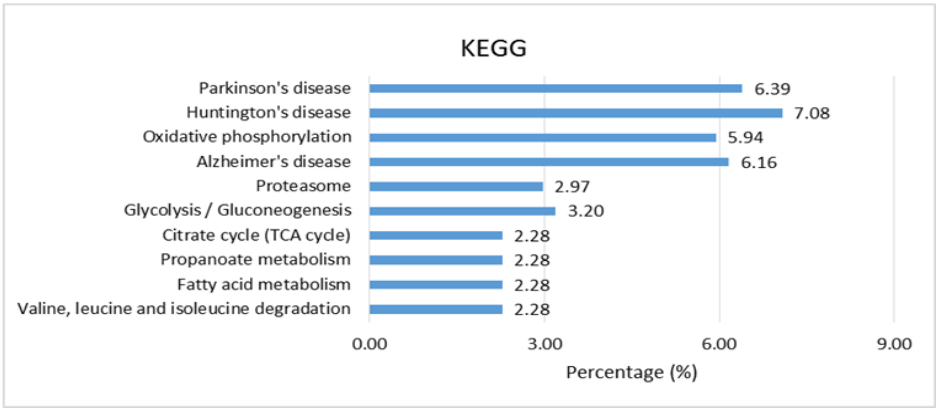


Figure 3.10.1: Top representative Gene Ontology (GO) and KEGG pathways enriched in CD patients as compared to control subjects. Functional annotation tool (FAT) of the differentially expressed proteins; the 10 most significantly enriched functional groups (GO terms) are shown ($p<10^{-13}$); BP: biological processes (panel **A**); MF: molecular functions (panel **B**) and KEGG pathways (panel **C**)). All classifications were significantly enriched compared to the entire proteomic dataset with $p< 0.05$ (Fisher's exact test).

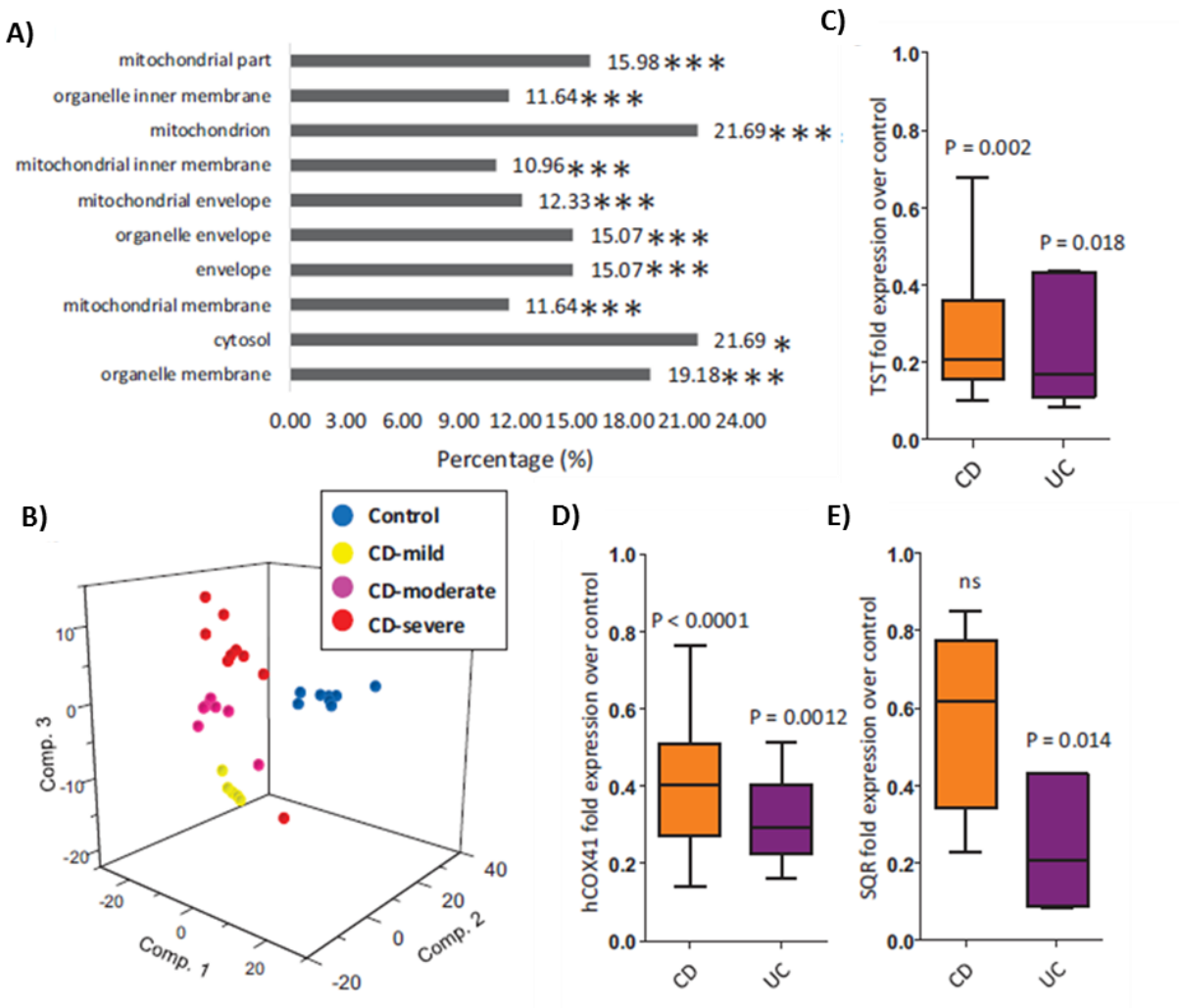


Figure 3.10.2: PLS-DA analysis of the host proteome identified the sulfide detoxification pathway as a functional biomarker of Crohn's disease. A, Functional annotation ("cellular component") analysis of the differentially expressed proteins; the 10 most significantly enriched functional groups (GO terms) are shown ($p<10^{-13}$). Asterisks denote classifications that were significantly enriched compared to the entire proteomic dataset. *$p<0.05$ and ***$p< 0.001$ (Fisher's exact test). B, PLS-DA of CD patients as a function of disease activity and controls (10 fold external validation of the model was performed on separate holdout validation sets of 5 randomly selected samples showing prediction accuracy of 75%). C-E, qRT-PCR of TST; cytochrome c oxidase subunit IV (hCOX41); sulfide dehydrogenase (SQR); normalized to control, n=5 for UC, 13 to 15 for CD and 10 to 15 for controls. Statistical significance was assessed using a two-tailed Mann-Whitney test. ns; not significant.

detoxification proteins and that the decreased abundance of these proteins is a hallmark of CD, and possibly UC, disease activity.

CHAPTER 4. Discussion

Our study represents a comprehensive, high-throughput survey of the gut microbiota and the host colonic proteome in pediatric patients with inflammatory bowel disease (IBD) at the time of disease diagnosis. We identify key microbial shifts that define the luminal-mucosal dysbiosis that occurs during first onset pediatric IBD. Prior studies using cohorts of patients post-treatment have linked microbial dysbiosis to IBD, however treatment intervention itself has been proven to alter the composition of gut microbiota [269]. Therefore, the identification of microbial taxa that potentially correlate with the disease etiology and/or progression requires newly diagnosed patients prior to the start of treatment. In addition to this unique cohort, our current study introduces a new microbiota sampling approach through mucosal aspirates. Additionally, we describe important differences between the microbial communities identified from this approach versus mucosal biopsies and/or stool samples. Finally, our work is the first of its kind to define the functional interplay between the host proteome and the gut microbiota in newly diagnosed early onset IBD**.**

The anatomic distribution and clinical presentation of IBD in children are different from adults, which indicate that pediatric IBD may constitute a unique phenotype of the disease. For example, IBD diagnosed at a young age mainly affects the large bowel [4, 5]. Furthermore, early onset IBD is characterized by a more severe clinical presentation and rapid progression of the disease within a short period of time as compared to adults with IBD [4-7]. Notably, pediatric cohorts exhibit fewer variables that affect the composition of the gut microbiota such as smoking, alcoholism and comorbidity as compared to adults [66, 270, 271]. Therefore, children constitute an ideal and valuable cohort for investigating the role of

gut microbiome in IBD. Some of the strongest evidence regarding the role of gut microbiota in IBD comes from genome wide association studies (GWAS). For instance, an investigation of 75,000 IBD cases and controls through GWAS identified 163 susceptibility loci associated with IBD [50]. The majority of these loci are involved in the immune pathways essential to maintaining the intestinal homeostasis [50]. In addition, the environmental risk factors associated with IBD, such as high fat diet, smoking, treatment with antibiotics, and sanitary conditions, have a direct effect on the composition of the gut microbiota [58, 66, 68]. Additionally, the association between dysbiosis within the gut microbiome and IBD has been reported by many microbiota profiling studies [94, 139]. Consequently, IBD may result from an aberrant immune response to one or more component(s) of the gut microbiota and/or their metabolic products.

Stool samples and mucosal biopsies are the two primary biological surrogates used to characterize the composition of gut microbiota. Fecal materials are easy to collect as they are non-invasive samples. However, they do not accurately reflect the composition of the resident mucosal microbiota [139]. For example, studies on healthy adults showed that stools harbor more Firmicutes and less Bacteroidetes as compared to the mucosal microbiota [91]. Additionally, Gevers *et al.* reported that the fecal microbiota from children with CD have fewer aerobes (Proteobacteria) and more anaerobes (*Bacteroides* and Clostridiales) as compared to mucosal biopsies from both the rectum and terminal ileum [139]. The mucosal microbiota are more health relevant as compared to the luminal microbiota since they are in direct contact with the intestinal epithelium and thus thought to be the key modulators of the immune system reactivity. Although mucosal biopsies are widely used to collect mucosal communities, they are an invasive and traumatic sampling method that generates more host

genetic material than microbial genetic material. In our study, we thought to collect mucosal-luminal microbiota through mucosal aspirates during colonoscopy procedures. Aspirate collection is less traumatic than biopsies and mainly consists of the loose mucus layer that harbors the majority of the mucosal microbiota (Figure 4.1) [152].

Initially, we aimed to compare the composition and diversity of the gut microbiota retrieved from stool samples, mucosal aspirates and mucosal biopsies from inflamed and non-inflamed tissues by sequencing the V6 region of the 16S rRNA gene. We first compared the microbial diversity between the three sample types in 4 children with CD. Although no significant difference in Chao1 indices was found, Shannon indices were significantly higher in stool samples as compared to biopsies from non-inflamed tissues. Moreover, the core microbiota identified from stool samples was the largest amongst the different sample types. In agreement with previous studies, our study reveals a significantly higher diversity in the microbial community of stool samples as compared to mucosal biopsies [91, 132, 139, 149, 272-278]. On the other hand, the microbial communities harbored by mucosal aspirates did not exhibit any significant difference in the diversity indices as compared to that of either biopsies or stool samples. Since mucosal aspirates collect the microbiota at the luminal-mucosal interface, it is more likely to exhibit a hybrid microbial diversity that lies in-between the mucosal biopsies and stool samples. Fecal microbiota, on the other hand, is composed of of free luminal microbiota in addition to the shed mucosal microbiota from varying intestinal regions [91]. Each intestinal region has its own microenvironment that shapes its specific microbial communities [279]. Therefore, it is more likely that stools harbor more diverse microbiota as compared to mucosal biopsies or aspirates which only collect the microbiota localized to a specific intestinal microenvironment.

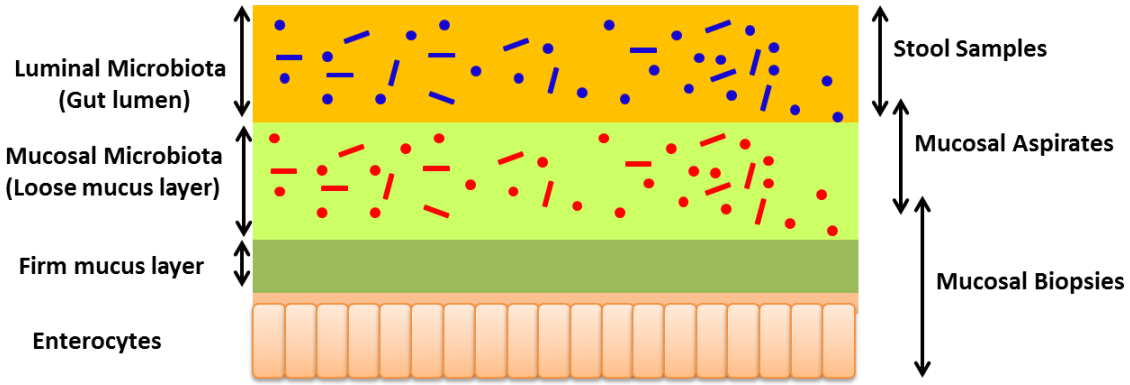


Figure 4.1: Schematic representation of the microbiota distribution across the gut lumen and different mucus layers of the large intestine. Mucosal biopsies mainly collect the mucosal microbiota, while stool samples collect the luminal microbiota. Mucosal aspirates collect the microbiota at the luminal mucosal interface.

The ANOSIM analysis comparing similarity of the gut microbiota between samples resulted in a higher R value when we categorized the samples by source individual as compared to sampling method (aspirates, biopsies or stools). This finding demonstrates that the interpersonal variabilities exhibit a higher contribution to the total microbial variability between samples relative to the sample type. These results coincide with Eckburg *et al.* who reported that differences between individuals explain the greatest variability among gut microbiota while the sampling approach explains most of the remaining variability [91]. Moreover, PCoA analyses showed that the microbiota identified from inflamed and non-inflamed biopsies of each individual paired together and clustered independently from the microbiota harbored by stool samples and mucosal aspirates. This result reveals a similarity between the mucosal microbiota identified from inflamed and non-inflamed colonic regions of CD subjects. While two studies reported differences in the mucosal microbiota between the inflamed non-inflamed regions of IBD [277, 280], the majority of studies have demonstrated that this is not the case [277, 281-285]. These two studies showed microbial differences only through the total bacterial load and fragments generated by terminal restriction fragment length polymorphisms (T-RFLP). However, they failed to identify any bacterial signatures that differentiate the two communities [277, 280]. Consequently, we may conclude through our analyses that the mucosal microbiota are host specific and uniformly distributed along the inflamed and non-inflamed areas of the right colon of CD patients. PCoA analyses also showed that the microbiota from the mucosal aspirates cluster closer to the fecal microbiota than to the microbiota from the mucosal biopsies which reveals that mucosal aspirates capture more luminal than mucosal microbiota.

The microbiota identified from CD stool samples and mucosal biopsies exhibit substantial differences including increased relative abundance of Firmicutes with a decreased relative abundance of Proteobacteria in the luminal, as compared to mucosal, microbiota. On the other hand, we did not observe these major differences when comparing the microbiota from mucosal aspirates to that of either stool samples or mucosal biopsies. Interestingly however, we did observe a gradual increase in the proportional abundance of the Firmicutes with a concomitant decrease of the Proteobacteria in mucosal biopsies, to mucosal aspirates, and to stool samples. This is consistent with many previous studies [66, 91, 139] which showed higher relative abundance of Firmicutes in stool samples relative to mucosal biopsies of different intestinal locations. However, a more recent study found no difference at the phyla level between rectal biopsies and fecal microbiota [286]. This discrepancy may be explained by differences in approach during biopsy collection, as Durban *et al*. [286] conducted their survey on un-prepped biopsies that could have been contaminated with fecal materials. Furthermore, the study of Durban *et al.* relied only on rectal biopsies while other studies utilized biopsies from different areas along the colon. Eckburg *et al.* showed that the mucosa associated microbiota from different colonic positions as well as the rectum differ from one another within the same individual, where the rectal mucosal microbiota is the most similar to the fecal microbiota compared to all other regions of the colon [91]. These findings reveal that the majority of the Firmicutes may reside at the lumen and that its relative abundance decreases as one gradually moves towards the colonic epithelium; in contrast, the opposite is true for Proteobacteria. This may be attributed to nutrient availability in the lumen such as undigested food fragments which can serve as a nutrient platform for Firmicutes [279, 287]. Moreover, host epithelial selection may have an impact on the pattern of microbial distribution across the gut width [288]. The studies of Morgan *et al*. and Gevers

et al. reported a depletion of Proteobacteria along with an enrichment of Firmicutes in the fecal microbiota as compared to the mucosal microbiota of IBD patients, which coincides with our observations for CD microbiota [66, 139]. It is known that CD inflammation results in a higher epithelial oxygen level, which shifts the gut microbiota towards more aerotolerant microbes [289]. The phylum Proteobacteria, which consists of aerobes and facultative anaerobes, is more resistant against inflammatory processes like oxidative stress as compared to the Firmicutes, which are obligate anaerobes [66, 289, 290]. Accordingly, we may assume that the Proteobacteria and not the Firmicutes may be better adapted to withstand the antimicrobial effects closer to the inflamed epithelium. We can equally assume that the distribution of the intestinal microbiome is not patchy, but instead exhibits a gradual continuous distribution across the width of the gut from the lumen and its contents, across the mucus layer and to the epithelium.

Our collective findings reveal that mucosal aspirates provide a distinct microbial profile compared to mucosal biopsies and stool samples. This microbial community exhibits an intermediary structure and diversity with similarities to the microbiota from both mucosal biopsies and stool samples. Moreover, it overcomes the disadvantages of the two more common sample types. Although the collection of colonoscopic mucosal aspirates is still an invasive approach, it is far less traumatic to the intestinal epithelium, covers a larger surface area and generates a higher ratio of microbe to host genetic material as compared to mucosal biopsies. The latter characteristic is the most important advantage for the detection of rare and less abundant microbes. In contrast to stool samples that provide a general view of the luminal microbiota, mucosal aspirates collect microbes that are in direct contact with the intestinal barrier which is more relevant to human health.

Several studies have characterized the gut microbiome of children under the age of 18 through stool samples [262-265]. As discussed before, stool samples harbor microbial communities that differ from resident mucosal microbes. Instead, we used mucosal aspirates to characterize the diversity and structure of the gut microbiota in healthy children between the ages of 3 to 18. We did not observe any difference in the diversity measures of the gut microbiota in healthy children along the length of the gut. These results concur with Eckburg *et al.* who reported no difference in the Shannon indices between different intestinal locations from cecum to rectum of three healthy individuals [91]. Recently, Gevers *et al*. showed the same results for new onset pediatric CD, where no significant difference in Chao1 was evident between the terminal ileum and rectum microbiota [139]. The absence of differential microbial diversity between the tested intestinal regions could be attributed to the high microbial variability between individuals. PCoA and ANOSIM similarity analyses show that the microbial communities from different intestinal regions of a given individual cluster together. These findings suggest that interpersonal variation in the biogeography of the human gut microbiome is greater than intrapersonal variation, which is in agreement with many previous studies conducted on healthy and IBD subjects [91, 94, 291]. However in contrast, Ahmed *et al*. showed that the mucosal microbiota from un-prepped ileal and colonic biopsies cluster based on the intestinal position and not by the host source [292]. This controversy may be explained by the absence of the bowel preparation prior to biopsy collection during this study.

Although we did not observe any differences in the microbial diversity in healthy children along the intestinal tract, the left and right colon of non-IBD children harbor a less diverse, smaller core microbiome as compared to their respective terminal ileum.

Furthermore, the microbiota composition in healthy children exhibits significant alterations from the terminal ileum to left colon. For example, Bacteroidetes (mainly the Bacteroidaceae and Porphyromondaceae families) are significantly decreased from the LC to the TI, whereas Firmicutes (primarily the lachnospiaraceae family) exhibit a significant enrichment from the LC to the RC. These results might be attributed to the growth requirements of these microbes, such as an optimal pH, oxygen requirements and nutrient availability. These conditions will vary significantly from the terminal ileum to the descending colon. For example, pH is known to be around 7.4 at the terminal ileum and drops suddenly to 5.5 at the cecum before it gradually increases as one moves from the cecum towards the rectum [178, 213]. Increasing the pH, by using streptomycin for example, has been associated with an enrichment of Bacteroidetes and a depletion of Firmicutes [293, 294]. This observation has recently been confirmed by Engevik *et al*., who reported an increase in luminal pH and Na^+ concentration is associated with increased relative abundance of Bacteroidetes and decreased relative abundance of Firmicutes [295]. The pH variability along the different regions of the gut will alter microbial metabolic activities respectively [178, 212]. While short-chain fatty acids, lactate and ethanol are produced at the cecum and ascending colon (mainly generated by the Firmicutes), the products of protein fermentation (e.g. ammonia and phenolic compounds) increase progressively towards the left colon [178, 212]. Accordingly, we can assume that the microbial composition at each intestinal site is a characteristic of the physiological status and physicochemical nature of that region.

As discussed previously in section 1.3, IBD may result from an aberrant immune response to one or more component(s) of the gut microbiota and/or their metabolic products. To investigate host-microbe mutualism in first onset pediatric IBD, we started with a

comparative study of the gut microbiota in 143 pediatric IBD patients (105 diagnosed cases of Crohn's disease and 38 with ulcerative colitis) against children with macroscopically and microscopically normal colons (control; n=74). All patients were newly diagnosed with IBD (inception cohort prior to the start of treatment). We collected the microbial community embedded within the mucus layer from 3 different intestinal locations: left colon, right colon and terminal ileum. Our findings reveal a comparable diversity of IBD microbiota to that of healthy individuals, which disagrees with many previous studies that reported decreased microbial diversity at the mucosa of adult IBD patients [94, 127, 128]. On the other hand, our results concur partially with Hansen and coauthors [140] who reported a similar microbial diversity between pediatric UC and control children, while CD patients exhibited a reduced microbiota diversity as compared to non-IBD control. Recently, Gevers *et al*. have confirmed our results by revealing a similar microbial diversity between pediatric CD and non-IBD controls [139]. Both of the aforementioned studies examined pediatric IBD cohorts of patients at the time of diagnosis which reveals that the similar microbial diversities that we observed between IBD and control children may be a characteristic of the new-onset pediatric IBD cohort.

Our results show microbial dysbiosis in both CD and UC at the three tested intestinal locations (LC, RC and TI). In agreement with prior literature, we identified Proteobacteria, Gammaproteobacteria, Betaproteobacteria, *Haemophilus*, Burkholderiales, Veillonellaceae, *Veillonella*, Neisseriacea, *Neisseria*, *Parvimonas*, *Gemella*, *Bacteroides spp.*, *Parabacteroides distasonis* and Gemellaceae as potential biomarkers of pediatric patients with CD [139, 296-299]. The majority of these microbes are oral and upper respiratory tract commensals [300]. However, they are known to colonize the inflamed mucosa [139]. This distribution may be attributable to their aerotolerant ability. For example, Proteobacteria

members, such as Enterobacteriaceae, and *Haemophilus* spp. are known to tolerate oxidative stress environments [66, 290]. Proteobacteria are known to synthesize glutathione which enables the bacteria to maintain their homeostasis during oxidative and acid stress [66]. Also, *Haemophilus influenza*, a commensal of the human upper respiratory tract, has evolved many mechanisms to survive and even intensify the oxidative stress during airway infection [290]. Another aerotolerant CD biomarker is *Veillonella spp.* such as *V. parvula* which has been demonstrated to resist H_2O_2 and even mitigates the H_2O_2 effect on other cocultured microbes [301]. These identified CD biomarkers in turn have the ability to exacerbate the inflammation through several ways. First, some bacteria have strong adhesive and invasive abilities for the epithelial cells such as *Gammaproteobacteria* members and the *Bacteroides/Prevotella* cluster [302, 303]. Second, Gram negative bacteria such as *E. coli*, *Haemophilus influenza* and *Veillonella parvula* produce lipopolysaccharides that can stimulate dendritic cells to release proinflammatory cytokines such as IL-12p70, TNF and IL-6 through the induction of toll like receptor 4 (TLR4) [304-306]. Finally, some of these microbes are potent H_2S producers through the fermentation of amino acids such as *Veillonella*, *Gemella*, *Parvimonas*, *Neisseria* and *E. coli* [176, 307-309]. Higher hydrogen sulfide generation, as a result of gut microbiota dysbiosis, is considered a potential model associated with the pathogenesis of IBD [71, 197]. This model assumes that an increased luminal concentration of H_2S disrupts the gut epithelial barrier and consequently intensifies the immune responses to the gut pathobionts and toxins [197].

In contrast to what is commonly observed [139, 228, 310], our results showed a higher relative abundance of *Faecalibacterium* in the CD microbiota as compared to the control group. Our finding agrees with Hansen *et al.* [140] who reported a higher incidence of *F. prausnitzii* in the mucosal microbiota of newly diagnosed pediatric CD but not in UC

patients. In contrast, Gevers *et al.* showed that *F. prausnitzii* exhibits a decreased relative abundance in the mucosal microbiota of children with first onset CD as compared to non-IBD children [139]. These two studies characterized the gut microbiota in first onset pediatric CD. The discrepancies between their findings may be explained by the different intestinal locations targeted in these studies. While Hansen *et al*. collected the mucosal biopsies from the distal colon, Gevers *et al*. used terminal ileum and rectal biopsies [139, 140]. In agreement with both studies, our results showed increased relative abundance of *F. prausnitzii* in the microbiota of left and right colons but not in the terminal ileum microbiota for children with CD compared to non-IBD controls. Consequently, the enrichment of *Faecalibacterium* may be a characteristic of the colonic microbiota in first onset pediatric CD.

Structure based analyses of the gut microbiome of UC patients show that Actinobacteria, *Veillonella*, *Prevotella*, *Peptostreptococcus*, *Neisseria*, *Gemella*, *Leptotrichia*, *Haemophilus* and *Porphyromonas* are enriched in UC as compared to the control patients. These microbes are known oral microbiota [300] and their abundance have been correlated to many oral diseases and malodours [176, 307, 309, 311]. In agreement with our results, Brito *et al.* have demonstrated that *Prevotella* and *Peptostreptococcus* species are overrepresented in the subgingival microbiota from CD and UC patients with untreated periodontitis [297]. Moreover, fluorescent *in-situ* hybridization of colonic specimens from UC patients have identified *Prevotella*/*Bacteroides* cluster, Enterobacteriaceae, as well as sulphate reducing bacteria as the predominant bacteria in the mucosa of UC patients [303]. The majority of these microbes including *Actinomyces*, *Veillonella*, *Prevotella*, *Gemella*, *Porphyromonas*, *Neisseria*, *Peptostreptococcus*, and *Leptotrichia* are potent hydrogen sulfide producers through the metabolism of sulfur containing amino acids [176, 307, 309, 311]. In

concordance with those observations, Morgan and coauthors showed an increase in the microbial metabolism of cysteine (the sulfur containing amino acid) in both CD and UC cases as compared to non-IBD individuals [66]. Although the association between the increased microbial production of H_2S and UC has been extensively reported, the microbial source for this hydrogen sulfide is still controversial. Several previous studies have linked the luminal H_2S generation in UC to sulfate reducing bacteria [199, 200, 312, 313]. However, other studies have reported no difference in the abundance of SRB in healthy and colitic subjects [202, 203]. In agreement with the latter, we did not detect a differential abundance of sulfate reducing bacteria (SRB) between the microbiota of UC and control subjects. The discrepancies may be explained by the different sample types, methodologies and/or patient cohorts employed in these studies [200, 313]. Knowing that cysteine is the main substrate of glutathione biosynthesis, the known antioxidant tripeptide, the metabolism of sulfur containing amino acids may enable these microbes to tolerate the redox potential of the inflammatory environment [66]. Another potential biomarker of UC that may withstand the oxidative stress during inflammation is *Vitreoscilla*. *Vitreoscilla* haemoglobin is known to be an antioxidant and has been recently identified as a probiotic that can ameliorate oxidative damage in the colon [314]. This antioxidative machinery may help *Vitreoscilla* live in high oxidative stress environments such as the inflamed UC colon. In general, we could conclude that inflammation-based alteration of the intestinal physicochemical characteristics such as the oxidative stress and perturbed nutrient availability may select for the organisms that could tolerate these challenges.

The microbiota of CD and UC are characterized by a depletion of many beneficial microbes that are essential for gut homeostasis and intestinal barrier function. For example, CD microbiota exhibit decreased abundance of Firmicutes, Lachnospiraceae, Clostridiaceae,

Erysipelotrichaceae, Erysipelotrichi, Verrucomicrobiaceae, *Eubacterium*, *Roseburia*, *Dorea* and *Blautia*, consistent with the results of Gevers and coauthors [139]. As well, the microbiota of UC subjects show a decrease in the relative abundance of Lachnospiraceae, Clostridia, *Roseburia*, *E. rectale*, *Ruminococcus*, Verrucomicrobia, *Dorea*, *Selenomonas*, *Blautia*, *Holdemania*, Porphyromondaceae, *Parabacteroides* and *Subdoligranulum* as compared to the control microbiota. Most of these taxa are potent producers of SCFA, mainly butyrate, which is a major energy source for colonocytes and a regulator of the intestinal barrier function [210, 315]. Moreover, the *Blautia* species *B. producta* along with 16 species related to *Clostridium*, *Ruminococcus* and *Eubacterium* have recently been identified as suppressors of gut inflammation by inducing $CD4^{+}FOXP3^{+}$ regulatory T (T_{reg})-cells [316]. T_{reg} cells are essential in maintaining gut homeostasis as they suppress differentiation of T_H17 cells. Together, these components contribute to the primary inflammatory pathway involved in the pathogenesis of CD [32, 35, 43, 44].

The microbes conserved during the co-evolution of humans and their gut microbiota (generalists) are evolved to utilize complex and dynamic nutrient sources as compared with host specific microbes that are adapted to few dietery substrates [317]. Hence, generalists are anticipated to contribute to human health more than specialists. These generalists constitute the so-called "core microbiota" and have the potential to be diagnostic markers [145, 318]. Our results reveal a higher relative abundance of *Fusobacterium nucleatum*, the potent H_2S producing oral microbe, in the core microbiota of CD cases as compared to controls. *F. nucleatum* has been previously associated with adult IBD and has been shown to promote tumorigenesis in $Apc^{min/+}$ mice [245, 319, 320]. Furthermore, the core microbiota of UC subjects showed a higher relative abundance of *Veillonella parvula*, the potent H_2S

generating taxon. *Veillonella* has been identified as a dominant oral microbe in IBD subjects where its abundance is strongly correlated to the increased release of proinflammatory cytokines such as IL-1β [296]. In addition, the left colon core microbiota of CD and UC patients, harbour more *Gemella* and *Leptotrichia*, respectively, than healthy children. Both genera are major contributors to the dysbiosis of the oral microbiota in IBD [296]. In comparison with the core microbiota of the control group, the core microbiota along the intestinal tract of both CD and UC patients is also characterized by fewer beneficial microbes such as Lachnospiraceae members which are essential for mucosal barrier function as previously discussed as well as in prior literature [316]. While the core microbial diversity was not affected in IBD patients, the CD and UC microbial communities are characterized by a smaller core microbiota with lower richness as compared to controls. This decrease in core microbiota may indicate a loss of microbiota homeostasis, which could lead to the opening of ecological niches, enabling microbial expansion of other taxa and/or colonization of disease-specific microbes.

The observed dysbiosis of the gut microbiota in pediatric IBD may lead to impaired microbial metabolism. To better understand how this perturbation is functionally linked to inflammation, we predicted the functionality of the microbial communities in CD, UC and control groups using PICRUST [66]. This approach infers the functionalities of the microbial communities based on their phylogenetic survey and a reference genomes database with prediction accuracy of 80-90% for gut microbial communities [66, 139]. We observed that functional dysbiosis in CD and UC microbiota includes enrichment of genes involved in pathogenesis such as the bacterial secretion systems. Secretion systems are responsible for extracellular release of proteins including proteases, pectinases, cellulases, phospholipases,

lipases and enterotoxins [321]. These proteins are known as virulence factors for pathobionts like *P. aeruginosa*, *Burkholderia spp*, *Erwinia,* and *E.coli* [321]. These particular bacterial species were observed to increase in the microbiota of CD as compared to the control microbiota in our study. These virulence factors can destroy the intestinal lining tissues and contribute to intestinal barrier damage and cytokine production [321, 322]. We also observed a depletion of microbial basic metabolism (such as decreased glycolysis, carbohydrate metabolism, thiamine metabolism and lysine as well as pantothenate and CoA biosynthesis) which reveals that the inflammatory environment may select for auxotrophic pathobionts like Proteobacteria members which translates to the observed increase in the lipopolysaccharide biosynthetic pathway, the tricarboxylic citric acid (TCA) cycle and the glutathione metabolism [66, 139]. Such selection could lead to more tissue damage either through toxin secretion or stimulation of cytokine release which in turn exacerbates the functional dysbiosis. For example, we observed an increase in taurine and hypotaurine metabolism in CD subjects as compared to the control group. Taurine, a sulfur containing compound that is secreted by the host to chelate bile acids, is subjected to microbial degradation by taurine-pyruvate aminotransferase to generate sulfite that is converted to H_2S by the action of microbial dissimilatory sulfite reductase [323]. Induced taurine metabolism has been shown to promote the bloom of rare sulfite reducing pathobiont, *Bilophila wadsworthia*, which in turn potentiated the development of colitis in *Il-10*$^{-/-}$ mice [71]. Collectively, we could conclude that tissue damage and perturbed metabolite availability in the gut of IBD subjects may select for pathobionts like auxotrophic organisms which then can magnify the tissue damage and lead to further functional and structural alterations within the microbial community.

Together, our results from surveying the microbial structures and functionalities in CD and UC patients have revealed that gut inflammation can act as a selective pressure that favours oxidative stress-resistant and auxotrophic microbes. This selectivity leads to expansion of potential colitogenic microbes which can disrupt the intestinal barrier and potentiate inflammatory responses towards gut pathobionts and/or their toxins. Therefore, we can describe host-microbe interactions in IBD as a dynamic process with a cross-feeding effect between the host and microbes.

To further explore which gut microbes are potential modulators of IBD inflammation, we sought to identify taxa associated with disease activity. Severe inflammation of both CD and UC at different intestinal locations is associated with the depletion of major butyrate producers such as Clostridia, Lachnospiraceae, *Lachnospira*, *Eubacterium, Coprococcus*, *Clostridium* and/or *Bifidobacterium adolescentis*. These microbes exhibit a significant decline in their respective relative abundance as disease severity increases. This result is consistent with many previous studies [139, 324] that identified an association between the clinical activity of IBD and a depletion of Clostridia, *Clostridium, Ruminococcus,* and *Eubacterium* members. Additionally, severe UC inflammation is associated with an increased abundance of certain taxa, namely *Corynebacterium*, *Vagococcus*, *Abiotrophia*, *Neisseria*, *Enterococcus*, *Staphylococcus aureus*, *Macrococcus*, and *Fibrobacter* at LC, and Flavobacteriaceae, *Catonella*, *Fusobacterium* and *Enterobacter* at RC. *Neisseria* and *Fusobacterium* are oral microbes and potent H_2S producers associated with oral malodour [309] Furthermore, enrichment of *Enterococcus* spp and *Enterobacteria* with a concomitant decrease in *Clostridium*

butyricum, *Ruminococcus albus*, and *Eubacterium rectale* has been, previously, associated with the clinical activity of adult UC [324].

We also observed direct association between Carnobacteriaceae, *Granulicatella*, *Mogibacterium*, *Proprionibacterium*, Bacillaceae, *Pantoea*, *Allobaculum* and *Atopobium* and CD disease activity. The association between *Atopobium* and the severity of inflammation in first onset CD has been confirmed, recently, by Gevers and coauthors [139]. Interestingly, we identified the same *Atopobium* OTU (OTU#529659) as a potential severity biomarker at both the LC and RC of CD patients. This OTU was further classified as *Atopobium parvulum*, a potent H_2S producer implicated in halitosis [325, 326]. Importantly, the increased abundance of *A. parvulum* with severity was not observed in UC and therefore is not simply a consequence of clinical inflammation, but instead a possible proinflammatory microbe in CD.

To evaluate the colitogenic potential of *A. parvulum*, we utilized colitic-susceptible *Il-10*$^{-/-}$ mice [253, 268] under specific pathogen free (SPF) conditions. These mice develop colitis spontaneously via interactions with their endogenous microbiota. This interaction is mediated by IL-23 which is produced by antigen presenting cells (APC) like dendritic cells. IL-23 in turn stimulates Th-17 lymphocytes to produce proinflammatory cytokines such as IL-17 and IL-6 [327]. IL-23-Th-17 pathway is known as a key pathway in IBD inflammation where many of the risk loci associated with IBD are related to this pathway [53]. Moreover, it is also an essential pathway for the defence against gut microbes and intestinal homeostasis [328]. Hence, we believe that the *Il-10*$^{-/-}$ murine colitis model is the most suitable model to study the contribution of gut microbes to IBD colitis. Our results indicate that *A. parvulum*-colonized *Il-10*$^{-/-}$ mice developed cecal atrophy and colon length reduction. Colonoscopy

imaging also revealed mucosal erythema, friability and mucosal ulceration in *A. parvulum*-colonized *Il-10*$^{-/-}$ mice compared to the healthy mucosa observed in control mice. Histological assessment of the intestinal tract showed evidence of inflammation with crypt hyperplasia, ulcers, goblet cell depletion and immune cell infiltration observed in the cecum and the distal part of the colon of *A. parvulum*-associated *Il-10*$^{-/-}$ mice compared to uninfected control *Il-10*$^{-/-}$ mice. Accordingly, histological inflammation scores were significantly higher in *A. parvulum* associated *Il-10*$^{-/-}$ mice compared to uninfected mice. At the molecular level, the colon of *A. parvulum* infected *Il-10*$^{-/-}$ mice showed increased *Il-12* and *Il-1β* mRNA accumulation compared to uninfected *Il-10*$^{-/-}$ mice. Notably, Hart *et al* have detected more IL-12-producing dendritic cells (DCs) in the lamina proporia of CD patients relative to control subjects [29]. Furthermore, IL-1β can promote chronic intestinal inflammation through stimulation of both IL-17A innate lymphoid cells and CD4(+) Th17 cells to produce IL-17 [329]. Together, these findings reveal that *A. parvulum* induces pancolitis in a genetically susceptible mouse model of IBD. This inflammation is Th-1/Th-17–mediated which is more similar to CD phenotype than UC phenotype [25, 26].

To gain mechanistic insights into the consequences of IBD-associated microbes on disease severity, we conducted an unbiased, quantitative proteomic analysis of mucosal biopsies from CD subjects with various disease severity and controls. Our data identified mitochondrial proteins as a major discriminant feature between control and CD metaproteome and even between subjects with different disease severity. The majority of proteins depleted in the CD metaproteome are proteins involved in energy generation and ATP production. Notably, components of mitochondrial hydrogen sulfide detoxification complex [172](Figure 4.2) were found to be the main proteins associated with disease severity. These proteins, namely the sulfur dioxygenase (ETHE1), the thiosulfate sulfur

transferase (TST), and the components of complexes III and IV of the mitochondrial respiratory chain, were down regulated in CD patients compared to controls. In agreement with prior literature [330, 331], secondary validation by qRT-PCR confirmed the repression of the TST transcript in CD and UC patients. Moreover, the expression levels of the cytochrome c oxidase subunit IV and the sulfide dehydrogenase genes (SQR), that contribute to the detoxification of H_2S [332], were significantly down-regulated in CD and/or UC patients. These findings indicate that transcriptional regulation contributes to the observed change in protein abundance and that the decreased abundance of these H_2S-detoxification proteins is a hallmark of CD disease activity and possibly UC. Importantly, these results would explain the previously observed increase of fecal sulfide levels in IBD patients [198].

From the characterization of the gut microbiota and host proteomic analysis, we demonstrate an alteration of the balance between bacterial-derived H_2S production and host-mediated detoxification of H_2S at mucosal-luminal interface. H_2S is now recognized as an important mediator of many physiological and pathological processes and has been associated with IBD and colorectal cancer [172, 333]. Moreover, the microbial metabolism of sulfur containing amino acids such as cysteine and methionine is enriched in IBD [66]. Hydrogen sulfide is known to inhibit the binding of oxygen to cytochrome c oxidase which in turn inhibits cellular respiration [334]. This inhibition disrupts the generation of adenosine tri-phosphate (ATP). In addition, H_2S inhibits butyrate oxidation which is the main source of energy for colonocytes [335]. Therefore, a higher concentration of H_2S might severely impair cellular bioenergetics, and thus in turn could potentially induce colonocyte starvation and death, disrupt the epithelial barrier and lead to inflammation.

To test the causative role of H_2S production in colitis, we assessed whether an H_2S

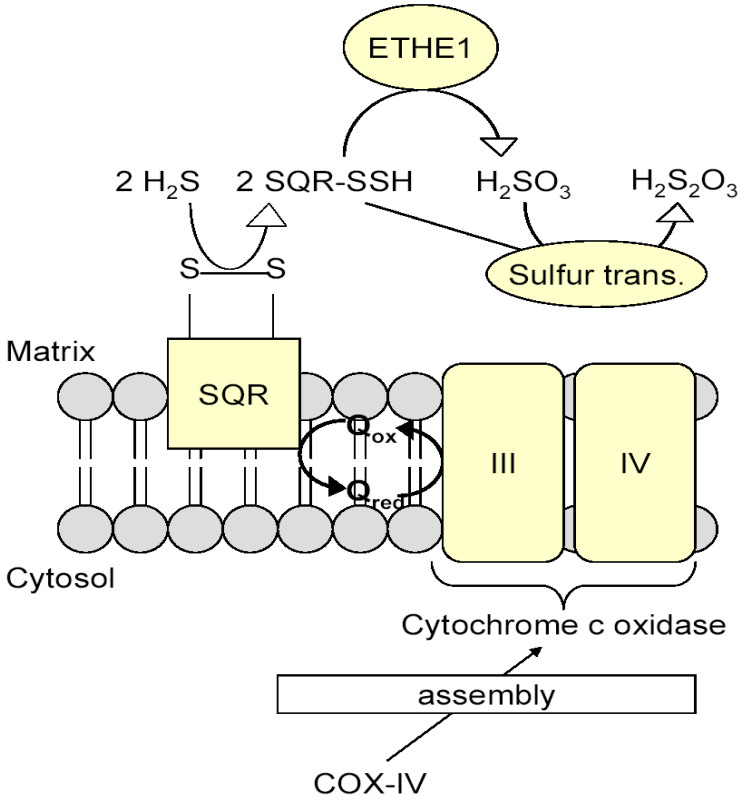


Figure 4.2: Model of mitochondrial H2S catabolism. The membrane bound sulfide dehydrogenase (SQR) oxidizes sulfide (H_2S) to persulfide (formed at one of the SQR's cysteines; SQR-SSH). The electrons are transferred to the mitochondrial respiratory chain (cytochrome c oxidase complex III and IV) via the quinone pool (Qox/Qred). The sulfur dioxygenase, ETHE1, oxidizes persulfides to sulfites (H_2SO_3) in the mitochondrial matrix. Finally, rhodanese (sulfur trans.) catalyzes the final reaction, which produces thiosulfite ($H_2S_2O_3$) by transferring a second persulfide from the SQR to sulfite. The cytochrome c oxidase subunit IV (COX-IV) is required for the assembly of the cytochrome c oxidase. Rhodanese comprises two iso-enzymes: thiosulfate sulfurtransferase (TST) and mercaptopyruvate sulfurtransferase (MST). Interestingly enough, the expression of both enzymes was previously shown to be lost in ulcerative colitis [209]. The figure was constructed by Dr. Alain Stintzi.

scavenger (bismuth) could alleviate *Atopobium*-induced colitis in *Il-10*$^{-/-}$ mice. Consistent with the first cohort, *Atopobium*-associated SPF mice developed severe colitis and exhibited significant increases in expression of pro-inflammatory cytokines. We also found that treatment with bismuth prevented *A. parvulum*-induced colitis. *Atopobium*-associated mice also exhibited an increased number of GALT (gut associated lymphoid tissue) foci as compared to non-associated mice. However bismuth treatment did not prevent GALT formation, indicating that *A. parvulum* induces GALT neogenesis in *Il-10*$^{-/-}$ mice via an H_2S-independent mechanism. It should be noted that the intestines of IBD patients also display a similarly increased number of lymphoid follicles[336]. Interestingly, elimination of GALT with LTßR-Ig treatment protects mice from developing colitis suggesting a role for GALT formation in the development of chronic intestinal inflammation [337]. Increased GALT foci in *Atopobium*-associated mice may lead to an aberrant expression of lymphoid adhesion molecules and unwanted T cell activation towards commensal microbes. To assess the role of these commensal microbes in colitis development, germ-free mice were mono-associated with *A. parvulum* and kept under gnotobiotic conditions. While these mice showed crypt hyperplasia and increased GALT foci, there were no signs of ulcerations, goblet cell depletion or immune cell infiltration. This result indicates that the gut microbiota may be required for the development of *Atopobium*-induced colitis. While bismuth treatment prevented GALT neogenesis in mice mono-associated with *A. parvulum*, the observed effect might not necessarily be due to H_2S scavenging but instead due to a potential antimicrobial activity of bismuth on *A. parvulum*, as evidenced by a reduced colonization level. The necessity of the gut microbiota to *A. parvulum*-induced colitis has been confirmed by using the DSS-induced colitis model. These mice showed significant reduction in colon length and weight, and increased IL-15 secretion as compared to *B. longum*-colonized mice. *Il-15* is an

NF-κB target gene that is essential for proliferation of intra-epithelial lymphocytes (IELs) in response to *NOD2* activation [338]. Accordingly, increased IL-15 in *Atopobium*-associated mice might exacerbate T cell immune responses against other commensal bacteria. Because both *A. parvulum* and gut microbiota are required for colitis development and because bismuth exhibits antimicrobial properties [339], we assessed the effect of bismuth on the gut microbiota composition of our SPF and *Atopobium*-associated mice. Our findings reveal a significant alteration in the microbial profile of *Atopobium*-associated mice as compared to SPF mice. Concomitantly to colitis prevention, bismuth administration altered the microbiota composition of these 2 groups of mice. Altogether, these results indicate that (1) *A. parvulum* colonization altered the composition of the gut microbiota (with a significant decrease in the abundance of the major butyrate-producers including *Eubacterium* and *Faecalibacterium* analogous to the microbiota composition of pediatric IBD patients); (2) the aberrant composition of the gut microbiota in *Atopobium*-associated mice is a major inducer of colitis; and (3) bismuth treatment restores the microbiota of these mice toward a healthier community (with an increased abundance of butyrate-producers).

4.2. Conclusions:

Our comparative assessment of the gut microbiota in CD, UC and non-IBD control children revealed several insights on the dysbiosis of this ecosystem in first onset pediatric IBD. The microbiota along the intestinal tract of IBD subjects lack the spatial heterogeneity observed in healthy subjects. While the microbial community of new-onset IBD patients exhibits a similar diversity to that of healthy controls, pediatric CD and UC subjects have a characteristic, depleted core microbiota relative to non-IBD children. IBD-specific microbial community is characterized by a loss of beneficial microbes, including inflammatory

suppressing taxa such as *Blautia producta* and Clostridiales. The loss of core and beneficial microbes may allow the expansion of opportunistic colonizers that favor inflammatory environments. For example, the identified IBD-enriched taxa such as Proteobacteria (e.g. Enterobacteriaceae), *Faecalibacterium prausnitzii*, and many oral microbes are aerotolerant microbes that can withstand the anaerobiosis shift and oxidative stress associated with IBD [66, 291, 340]. This hypothesis is supported by the study of Lupp *et al.* who reported that host mediated inflammation has been shown to alter the intestinal microbiota and promotes enrichment of Enterobacteriaceae [341]. This microbial shift may in turn exacerbate the intestinal inflammation. Our findings begin to shed light on the synergism between the depletion of butyrate producing bacteria and the enrichment of H_2S generating microbes as a potential pathogenic mechanism of IBD. IBD core microbiota are characterized by enrichment of potent H_2S producers such as *Veillonella dispar* and *Fusobacterium nucleatum* with a concomitant loss of butyrate producing bacteria. Moreover, the severity of colonic inflammation of CD and UC patients correlates with *Atopobium parvulum* and *Fusobacterium,* respectively. *A. parvulum, Fusobacterium nucleatum* and *Veillonella dispar* are known to produce H_2S through protein fermentation [342]. A higher concentration of H_2S has been proposed to be associated with the pathogenesis of IBD [71, 146, 198], though some studies reported that the luminal H_2S level is not elevated in UC [201]. This discrepancy may be explained by the approach used to quantify the H_2S level. For example, Levine *et al.* showed that H_2S release from incubated UC stool is 3-4 fold higher than normal feces over a period of 24 h [198], while Moore *et al.* reported that the stool sulfide measured using spectrophotometric approach on stool slurries is not elevated in UC as compared to control subjects [201]. While the elevated H_2S in IBD has long been linked to sulfate-reducing bacteria (SRB), several studies have failed to demonstrate an association between

SRB and IBD [202, 203, 333]. Instead, our study demonstrates a key role for microbes that produce H_2S through protein fermentation in CD pathogenesis. Butyrate is known to activate the expression of the genes encoding the host mitochondrial H_2S detoxification components [343]. Our proteomic analyses indicate a diminished capacity for H_2S detoxification by IBD patients. Therefore, the depletion of butyrate-producing microbes from the gut microbiota may disable the host H_2S defence systems. This "disarmed" host would be highly susceptible to further damage caused by enhanced H_2S production, resulting in metabolic stress and subsequently increased mucosal inflammation (Figure 4.3). Interestingly, variants in mitochondrial DNA, which result in increased metabolic activities, protect mice from colitis [344]. This is in agreement with the important role of the mitochondria in modulating the mucosal barrier. More recently, excess H_2S has been shown to act as an autocrine T-cell activator, potentially contributing to unwanted T-cell responses against commensal bacteria [345]; consistent with our observation that the gut microbiota is required for *A. parvulum*-induced experimental colitis. Given the essential role of butyrate in regulating regulatory T cell (Treg) homeostasis and the critical role of Treg in limiting intestinal inflammation [315] H_2S production may also interfere with this process by impairing butyrate oxidation, and thus might lead to increased colitis severity. This result emphasises the importance of the microbial community and its interaction with the host in the pathogenesis of IBD. Altogether, our study represents the first report that begins to shed light on the pathogenic mechanisms of early IBD onset.

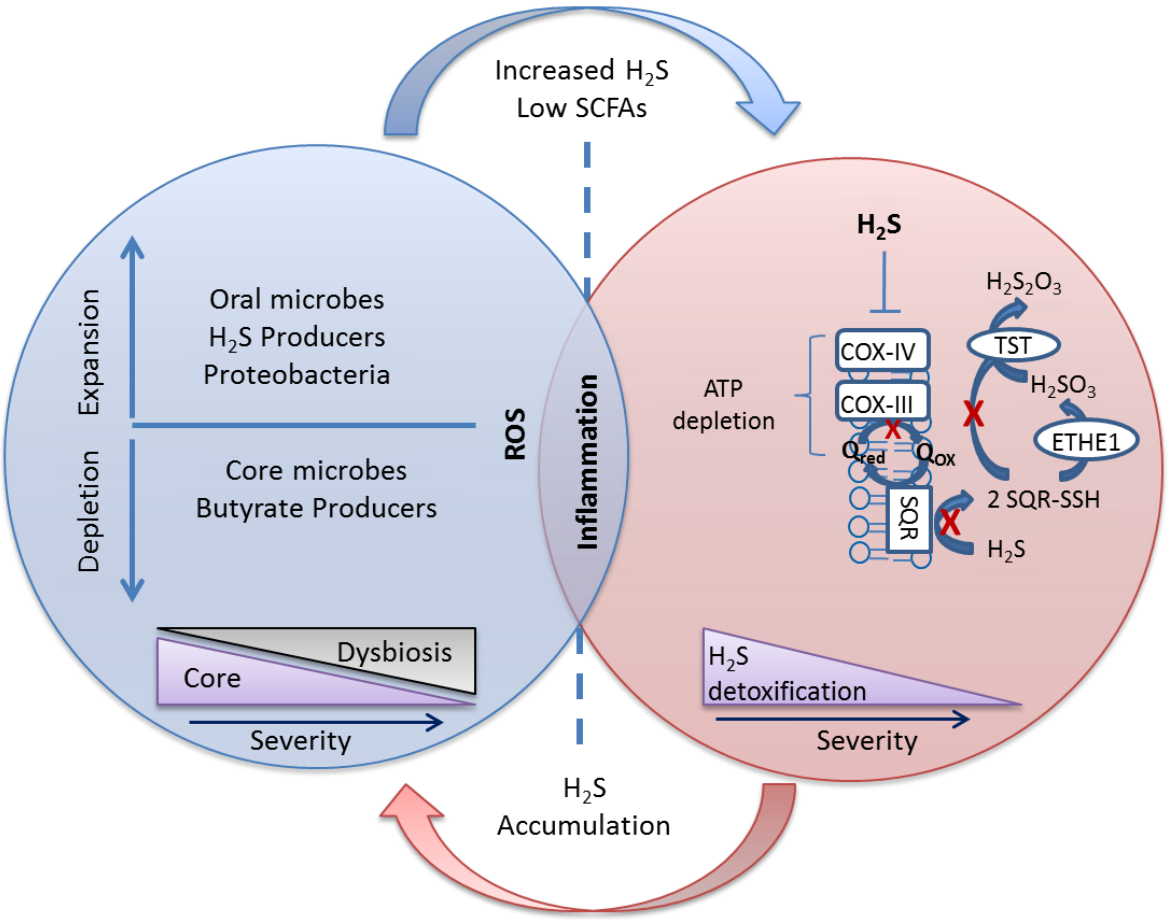


Figure 4.3: Proposed host-microbe interaction in pediatric IBD. Gut microbiota dysbiosis in pediatric IBD is characterized by loss of core microbes, including beneficial SCFA producers. In association with the inflammation selective pressure, this depletion of core microbes allows for the expansion of minor opportunistic bacteria such as oral microbes involved in hydrogen sulfide production and other inflammation-resistant bacteria like aerotolerant microbes. This altered microbiota might generate colitogenic factors such as a high concentration of hydrogen sulfide. Accordingly, hydrogen sulfide accumulation may inhibit the oxidative machinery of the colonocytes, which in association with depleted energy sources like SCFA reduces colonocyte energetic activity and induce colonocyte starvation, which in turn disables the host hydrogen sulfide detoxification machinery. Consequently, this metabolic stress may disrupt the epithelial barrier, induce unwanted T-responses against commensal microbes and potentiate the mucosal inflammation.

4.3. Future Directions.

The main findings of this study are consistent with altered microbiota-host mutualism and enrichment of H_2S generating microbes drive the severity of CD and possibly UC. However solid proof regarding the role of H_2S in the pathogenesis of early onset IBD is still lacking. Although, bismuth subsalicylate reduces the *A. parvulum*-induced inflammation, this effect was evident also in non *A. parvulum* associated mice. Moreover, the level of colonization of *A. parvulum* in bismuth-treated mice was significantly lower than untreated mice. Therefore, the contribution of bismuth to the reduced inflammation might arise from its antimicrobial activity rather than the H_2S scavenging ability. Further studies are required to confirm the role of H_2S in the pathogenesis and the severity of intestinal inflammation. To further prove that, an *A. parvulum* mutant strain that lacks H_2S production capability should be constructed. Inoculating this strain in the same colitis model employed here and comparing its results to the wild type strain would serve to clarify the implication of H_2S in IBD.

The role of H_2S-butyrate metabolism in early onset IBD relies mainly on the relative abundance of their bacterial producers and is strongly supported by the metaproteomic survey done on human colonic biopsies. However, the relative abundance of certain taxa may not accurately reflect the overall functional activity of the community. Increased luminal concentration of H_2S has been proposed to be associated with adult UC [198]. However, assessment of hydrogen sulfide concentration in CD and first onset IBD patients is still missing. To confirm the association between high hydrogen sulfide concentration and CD or early onset IBD, experiments to detect hydrogen sulfide concentration or its substrate metabolism in relation to the disease severity are required. This could be done via

metatranscriptomic analysis or qRT-PCR targeting key genes involved in cysteine-methionine metabolism. Furthermore, the diversity of H_2S producers and butyrate producers identified here was determined based on 16S rRNA data. This would limit our results to the known major producers only which results in inaccurate overall prediction of this functional groups diversity. By using butyryl-CoA: acetate-CoA- transferase-based 454 pyrosequencing, Vital *et al.* have shown more specific functional characterization of the gut butyrate producers [346]. Unpublished work from our lab has characterized the diversity of butyrate producers in pediatric CD, UC and control groups by 454-pyrosequencing of butyryl-CoA: acetate-CoA- transferase. The results from that work demonstrate a decreased relative abundance of the major butyrate producers in the microbiota of CD and UC as compared to the control group. In addition and in agreement with our results an increase in the relative abundance of *F. prausnitzii* was also observed in the microbiota of pediatric patients with CD as compared to the control subjects. Notably, this work has also identified unclassified clusters of bacteria that have not been recognized before as butyrate producers. Consequently, applying similar functional gene based sequencing may enable us to accurately characterize the diversity of other functional groups in IBD such as H_2S producers, from either organic or inorganic sulfur containing compounds.

Finally, further research is required to determine whether the identified dysbiosis in early onset IBD is a cause or a consequence of the inflammation; furthermore, what are the consequences of this altered community on disease severity? This could be revealed by monitoring gut microbiota variabilities over a long period of disease duration. We could start at the time of diagnosis and correlate the changes in microbiota composition over time to the disease severity. In addition to 16S rRNA characterization, metagenomic,

metatranscriptomic or metabolomics studies will be necessary to reveal the biomolecular mechanisms by which the microbiota is affected or contribute to the inflammation.

The work presented here represents the second largest survey of gut microbiota in first onset pediatric CD and the largest one for UC. Moreover, it is the first study that provides mechanistic insights into the role of microbiota in disease pathogenesis in a unique cohort of patients. We characterized the composition of gut microbiota along the intestinal tract of IBD patients and non-IBD control subjects. The microbial community of IBD patients is characterized by depleted core and beneficial microbes with a concomitant enrichment of Proteobacteria members, H_2S producing bacteria and oral microbes. Correlating this dysbiosis to disease severity has identified enrichment of H_2S producers in association with depletion of butyrate producers as severity biomarkers. Studies using colitis murine models and mass spectrometry-based assessment of human colonic proteome identified *A. parvulum* and impaired H_2S metabolism as major mediators of CD severity. Altogether, we can conclude that altered host-microbe mutualism and the bloom of *A. parvulum* are major drivers of disease severity in first onset pediatric IBD. However, future work is required to elucidate the biomolecular mechanisms by which the disruption of butyrate-H_2S metabolic axis could potentiate inflammation.

CHAPTER 5. References

1. Tremaine WJ: **Diagnosis and treatment of indeterminate colitis**. *Gastroenterology & hepatology* 2011, **7**(12):826-828.
2. Baumgart DC, Sandborn WJ: **Inflammatory bowel disease: clinical aspects and established and evolving therapies**. *Lancet* 2007, **369**(9573):1641-1657.
3. Levine JS, Burakoff R: **Extraintestinal manifestations of inflammatory bowel disease**. *Gastroenterology & hepatology* 2011, **7**(4):235-241.
4. Van Limbergen J, Russell RK, Drummond HE, Aldhous MC, Round NK, Nimmo ER, Smith L, Gillett PM, McGrogan P, Weaver LT *et al*: **Definition of phenotypic characteristics of childhood-onset inflammatory bowel disease**. *Gastroenterology* 2008, **135**(4):1114-1122.
5. Nieuwenhuis EE, Escher JC: **Early onset IBD: what's the difference?** *Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver* 2008, **40**(1):12-15.
6. Witte J, Shivananda S, Lennard-Jones JE, Beltrami M, Politi P, Bonanomi A, Tsianos EV, Mouzas I, Schulz TB, Monteiro E *et al*: **Disease outcome in inflammatory bowel disease: mortality, morbidity and therapeutic management of a 796-person inception cohort in the European Collaborative Study on Inflammatory Bowel Disease (EC-IBD)**. *Scand J Gastroenterol* 2000, **35**(12):1272-1277.
7. Israeli E, Ryan JD, Shafer LA, Bernstein CN: **Younger Age at Diagnosis Is Associated With Panenteric, but Not More Aggressive, Crohn's Disease**. *Clin Gastroenterol H* 2014, **12**(1):72-U117.
8. Escher JC, Dias JA, Bochenek K, Buderus S, de Mesquita MB, Bujanover Y, Buller HA, Chong SKF, Cucchiara S, Fell JME *et al*: **Inflammatory bowel disease in children and adolescents: Recommendations for diagnosis - The Porto criteria**. *J Pediatr Gastr Nutr* 2005, **41**(1):1-7.
9. Motil KJ, Grand RJ, Davis-Kraft L, Ferlic LL, Smith EO: **Growth failure in children with inflammatory bowel disease: a prospective study**. *Gastroenterology* 1993, **105**(3):681-691.
10. Molodecky NA, Soon IS, Rabi DM, Ghali WA, Ferris M, Chernoff G, Benchimol EI, Panaccione R, Ghosh S, Barkema HW *et al*: **Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review**. *Gastroenterology* 2012, **142**(1):46-54 e42; quiz e30.
11. Burisch J, Munkholm P: **Inflammatory bowel disease epidemiology**. *Current opinion in gastroenterology* 2013.
12. Benchimol EI, Manuel DG, Guttmann A, Nguyen GC, Mojaverian N, Quach P, Mack DR: **Changing age demographics of inflammatory bowel disease in Ontario, Canada: a population-based cohort study of epidemiology trends**. *Inflammatory bowel diseases* 2014, **20**(10):1761-1769.
13. Lowe AM, Roy PO, M BP, Michel P, Bitton A, St-Onge L, Brassard P: **Epidemiology of Crohn's disease in Quebec, Canada**. *Inflammatory bowel diseases* 2009, **15**(3):429-435.

14. Bernstein CN, Wajda A, Svenson LW, MacKenzie A, Koehoorn M, Jackson M, Fedorak R, Israel D, Blanchard JF: **The epidemiology of inflammatory bowel disease in Canada: a population-based study**. *The American journal of gastroenterology* 2006, **101**(7):1559-1568.
15. Wilson J, Hair C, Knight R, Catto-Smith A, Bell S, Kamm M, Desmond P, McNeil J, Connell W: **High incidence of inflammatory bowel disease in Australia: a prospective population-based Australian incidence study**. *Inflammatory bowel diseases* 2010, **16**(9):1550-1556.
16. Rocchi A, Benchimol EI, Bernstein CN, Bitton A, Feagan B, Panaccione R, Glasgow KW, Fernandes A, Ghosh S: **Inflammatory bowel disease: a Canadian burden of illness review**. *Canadian journal of gastroenterology = Journal canadien de gastroenterologie* 2012, **26**(11):811-817.
17. Benchimol EI, Guttmann A, Griffiths AM, Rabeneck L, Mack DR, Brill H, Howard J, Guan J, To T: **Increasing incidence of paediatric inflammatory bowel disease in Ontario, Canada: evidence from health administrative data**. *Gut* 2009, **58**(11):1490-1497.
18. Sonnenberg A: **Age distribution of IBD hospitalization**. *Inflammatory bowel diseases* 2010, **16**(3):452-457.
19. Weinstein TA, Levine M, Pettei MJ, Gold DM, Kessler BH, Levine JJ: **Age and family history at presentation of pediatric inflammatory bowel disease**. *J Pediatr Gastr Nutr* 2003, **37**(5):609-613.
20. Orholm M, Munkholm P, Langholz E, Nielsen OH, Sorensen TI, Binder V: **Familial occurrence of inflammatory bowel disease**. *The New England journal of medicine* 1991, **324**(2):84-88.
21. De Hertogh G, Aerssens J, Geboes KP, Geboes K: **Evidence for the involvement of infectious agents in the pathogenesis of Crohn's disease**. *World J Gastroenterol* 2008, **14**(6):845-852.
22. Shih DQ, Targan SR, McGovern D: **Recent advances in IBD pathogenesis: genetics and immunobiology**. *Curr Gastroenterol Rep* 2008, **10**(6):568-575.
23. Asakura H, Suzuki K, Honma T: **Recent advances in basic and clinical aspects of inflammatory bowel disease: which steps in the mucosal inflammation should we block for the treatment of inflammatory bowel disease?** *World J Gastroenterol* 2007, **13**(15):2145-2149.
24. Round JL, Mazmanian SK: **The gut microbiota shapes intestinal immune responses during health and disease**. *Nature reviews Immunology* 2009, **9**(5):313-323.
25. Fuss IJ, Neurath M, Boirivant M, Klein JS, de la Motte C, Strong SA, Fiocchi C, Strober W: **Disparate CD4+ lamina propria (LP) lymphokine secretion profiles in inflammatory bowel disease. Crohn's disease LP cells manifest increased secretion of IFN-gamma, whereas ulcerative colitis LP cells manifest increased secretion of IL-5**. *Journal of immunology* 1996, **157**(3):1261-1270.
26. Sartor RB: **Mechanisms of disease: pathogenesis of Crohn's disease and ulcerative colitis**. *Nature clinical practice Gastroenterology & hepatology* 2006, **3**(7):390-407.

27. Matsuoka K, Inoue N, Sato T: **T-bet upregulation and subsequent interleukin 12 stimulation are essential for induction of Th1 mediated immunopathology in Crohn's disease (vol 53, pg 1303, 2004)**. *Gut* 2004, **53**(11):1722-1722.
28. Neurath MF, Fuss I, Kelsall BL, Stuber E, Strober W: **Antibodies to Interleukin-12 Abrogate Established Experimental Colitis in Mice**. *J Exp Med* 1995, **182**(5):1281-1290.
29. Hart AL, Al-Hassi HO, Rigby RJ, Bell SJ, Emmanuel AV, Knight SC, Kamm MA, Stagg AJ: **Characteristics of intestinal dendritic cells in inflammatory bowel diseases**. *Gastroenterology* 2005, **129**(1):50-65.
30. Targan SR, Hanauer SB, van Deventer SJ, Mayer L, Present DH, Braakman T, DeWoody KL, Schaible TF, Rutgeerts PJ: **A short-term study of chimeric monoclonal antibody cA2 to tumor necrosis factor alpha for Crohn's disease. Crohn's Disease cA2 Study Group**. *The New England journal of medicine* 1997, **337**(15):1029-1035.
31. Neurath MF, Fuss I, Kelsall BL, Stuber E, Strober W: **Antibodies to interleukin 12 abrogate established experimental colitis in mice**. *J Exp Med* 1995, **182**(5):1281-1290.
32. Kobayashi T, Okamoto S, Hisamatsu T, Kamada N, Chinen H, Saito R, Kitazume MT, Nakazawa A, Sugita A, Koganei K *et al*: **IL23 differentially regulates the Th1/Th17 balance in ulcerative colitis and Crohn's disease**. *Gut* 2008, **57**(12):1682-1689.
33. Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ, Steinhart AH, Abraham C, Regueiro M, Griffiths A *et al*: **A genome-wide association study identifies IL23R as an inflammatory bowel disease gene**. *Science* 2006, **314**(5804):1461-1463.
34. Kullberg MC: **IL-23 plays a key role in Helicobacter hepaticus-induced T cell-dependent colitis**. *Immunology* 2007, **120**:25-26.
35. Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ, Steinhart AH, Abraham C, Regueiro M, Griffiths A *et al*: **A genome-wide association study identifies IL23R as an inflammatory bowel disease gene**. *Science* 2006, **314**(5804):1461-1463.
36. Elson CO, Cong YZ, Weaver CT, Schoeb TR, McClanahan TK, Fick RB, Kastelein RA: **Monoclonal anti-interleukin 23 reverses active colitis in a T cell-mediated model in mice**. *Gastroenterology* 2007, **132**(7):2359-2370.
37. Cox JH, Kljavin NM, Ota N, Leonard J, Roose-Girma M, Diehl L, Ouyang W, Ghilardi N: **Opposing consequences of IL-23 signaling mediated by innate and adaptive cells in chemically induced colitis in mice**. *Mucosal Immunol* 2012, **5**(1):99-109.
38. Strober W, Fuss IJ: **Proinflammatory Cytokines in the Pathogenesis of Inflammatory Bowel Diseases**. *Gastroenterology* 2011, **140**(6):1756-U1782.
39. Ghoreschi K, Laurence A, Yang XP, Tato CM, McGeachy MJ, Konkel JE, Ramos HL, Wei L, Davidson TS, Bouladoux N *et al*: **Generation of pathogenic T(H)17 cells in the absence of TGF-beta signalling**. *Nature* 2010, **467**(7318):967-U144.
40. O'Connor W, Kamanaka M, Booth CJ, Town T, Nakae S, Iwakura Y, Kolls JK, Flavell RA: **A protective function for interleukin 17A in T cell-mediated intestinal inflammation**. *Nat Immunol* 2009, **10**(6):603-U665.

41. Zenewicz LA, Yancopoulos GD, Valenzuela DM, Murphy AJ, Stevens S, Flavell RA: **Innate and Adaptive Interleukin-22 Protects Mice from Inflammatory Bowel Disease**. *Immunity* 2008, **29**(6):947-957.
42. Izcue A, Hue S, Buonocore S, Arancibia-Carcamo CV, Ahern PP, Iwakura Y, Maloy KJ, Powrie F: **Interleukin-23 restrains regulatory T cell activity to drive T cell-dependent colitis**. *Immunity* 2008, **28**(4):559-570.
43. Ivanov, II, Frutos Rde L, Manel N, Yoshinaga K, Rifkin DB, Sartor RB, Finlay BB, Littman DR: **Specific microbiota direct the differentiation of IL-17-producing T-helper cells in the mucosa of the small intestine**. *Cell host & microbe* 2008, **4**(4):337-349.
44. Powrie F, Maloy KJ: **Immunology. Regulating the regulators**. *Science* 2003, **299**(5609):1030-1031.
45. Fuss IJ, Heller F, Boirivant M, Leon F, Yoshida M, Fichtner-Feigl S, Yang ZQ, Exley M, Kitani A, Blumberg RS *et al*: **Nonclassical CD1d-restricted NK T cells that produce IL-13 characterize an atypical Th2 response in ulcerative colitis**. *J Clin Invest* 2004, **113**(10):1490-1497.
46. Van Limbergen J, Wilson DC, Satsangi J: **The genetics of Crohn's disease**. *Annu Rev Genomics Hum Genet* 2009, **10**:89-116.
47. Halme L, Paavola-Sakki P, Turunen U, Lappalainen M, Farkkila M, Kontula K: **Family and twin studies in inflammatory bowel disease**. *World J Gastroenterol* 2006, **12**(23):3668-3672.
48. Imielinski M, Baldassano RN, Griffiths A, Russell RK, Annese V, Dubinsky M, Kugathasan S, Bradfield JP, Walters TD, Sleiman P *et al*: **Common variants at five new loci associated with early-onset inflammatory bowel disease**. *Nat Genet* 2009, **41**(12):1335-1340.
49. Kugathasan S, Baldassano RN, Bradfield JP, Sleiman PM, Imielinski M, Guthery SL, Cucchiara S, Kim CE, Frackelton EC, Annaiah K *et al*: **Loci on 20q13 and 21q22 are associated with pediatric-onset inflammatory bowel disease**. *Nat Genet* 2008, **40**(10):1211-1215.
50. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, Lee JC, Schumm LP, Sharma Y, Anderson CA *et al*: **Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease**. *Nature* 2012, **491**(7422):119-124.
51. Christodoulou K, Wiskin AE, Gibson J, Tapper W, Willis C, Afzal NA, Upstill-Goddard R, Holloway JW, Simpson MA, Beattie RM *et al*: **Next generation exome sequencing of paediatric inflammatory bowel disease patients identifies rare and novel variants in candidate genes**. *Gut* 2012.
52. Mao H, Yang W, Lee PP, Ho MH, Yang J, Zeng S, Chong CY, Lee TL, Tu W, Lau YL: **Exome sequencing identifies novel compound heterozygous mutations of IL-10 receptor 1 in neonatal-onset Crohn's disease**. *Genes Immun* 2012.
53. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, Lee JC, Schumm LP, Sharma Y, Anderson CA *et al*: **Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease**. *Nature* 2012, **491**(7422):119-124.

54. Hamilton MJ, Snapper SB, Blumberg RS: **Update on biologic pathways in inflammatory bowel disease and their therapeutic relevance**. *J Gastroenterol* 2012, **47**(1):1-8.
55. Ananthakrishnan AN, Higuchi LM, Huang ES, Khalili H, Richter JM, Fuchs CS, Chan AT: **Aspirin, nonsteroidal anti-inflammatory drug use, and risk for Crohn disease and ulcerative colitis: a cohort study**. *Ann Intern Med* 2012, **156**(5):350-359.
56. Bernstein CN, Shanahan F: **Disorders of a modern lifestyle: reconciling the epidemiology of inflammatory bowel diseases**. *Gut* 2008, **57**(9):1185-1191.
57. Cosnes J: **What is the link between the use of tobacco and IBD?** *Inflammatory bowel diseases* 2008, **14 Suppl 2**:S14-15.
58. Hou JK, Abraham B, El-Serag H: **Dietary intake and risk of developing inflammatory bowel disease: a systematic review of the literature**. *The American journal of gastroenterology* 2011, **106**(4):563-573.
59. Leslie WD, Miller N, Rogala L, Bernstein CN: **Vitamin D status and bone density in recently diagnosed inflammatory bowel disease: the Manitoba IBD Cohort Study**. *The American journal of gastroenterology* 2008, **103**(6):1451-1459.
60. Shaw SY, Blanchard JF, Bernstein CN: **Association between the use of antibiotics in the first year of life and pediatric inflammatory bowel disease**. *The American journal of gastroenterology* 2010, **105**(12):2687-2692.
61. Mahid SS, Minor KS, Soto RE, Hornung CA, Galandiuk S: **Smoking and inflammatory bowel disease: a meta-analysis**. *Mayo Clinic proceedings* 2006, **81**(11):1462-1471.
62. Timmer A SL, Martin F: **Oral contraceptive use and smoking are risk factors for relapse in Crohn's disease. The Canadian Mesalamine for Remission of Crohn's Disease Study Group.** *Gastroenterology* 1998, **114**(6):1143-1150.
63. Cosnes J, Beaugerie L, Carbonnel F, Gendre JP: **Smoking cessation and the course of Crohn's disease: an intervention study**. *Gastroenterology* 2001, **120**(5):1093-1099.
64. Boyko EJ, Perera DR, Koepsell TD, Keane EM, Inui TS: **Effects of cigarette smoking on the clinical course of ulcerative colitis**. *Scand J Gastroenterol* 1988, **23**(9):1147-1152.
65. Bastida G, Beltran B: **Ulcerative colitis in smokers, non-smokers and ex-smokers**. *World J Gastroenterol* 2011, **17**(22):2740-2747.
66. Morgan XC, Tickle TL, Sokol H, Gevers D, Devaney KL, Ward DV, Reyes JA, Shah SA, LeLeiko N, Snapper SB *et al*: **Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment**. *Genome Biol* 2012, **13**(9):R79.
67. Wang H, Zhao JX, Hu N, Ren J, Du M, Zhu MJ: **Side-stream smoking reduces intestinal inflammation and increases expression of tight junction proteins**. *World J Gastroenterol* 2012, **18**(18):2180-2187.
68. De Filippo C, Cavalieri D, Di Paola M, Ramazzotti M, Poullet JB, Massart S, Collini S, Pieraccini G, Lionetti P: **Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa**. *Proc Natl Acad Sci U S A* 2010, **107**(33):14691-14696.
69. Martinez I, Lattimer JM, Hubach KL, Case JA, Yang JY, Weber CG, Louk JA, Rose DJ, Kyureghian G, Peterson DA *et al*: **Gut microbiome composition is linked to whole grain-induced immunological improvements**. *Isme Journal* 2013, **7**(2):269-280.

70. Smith MI, Yatsunenko T, Manary MJ, Trehan I, Mkakosya R, Cheng J, Kau AL, Rich SS, Concannon P, Mychaleckyj JC *et al*: **Gut microbiomes of Malawian twin pairs discordant for kwashiorkor**. *Science* 2013, **339**(6119):548-554.
71. Devkota S, Wang YW, Leone V, Musch M, Nadimpalli A, Antonopoulos D, Jabri B, Chang E: **Dietary fat-induced taurocholic acid production promotes pathobiont and colitis in IL-10-/- mice**. *Faseb J* 2012, **26**.
72. Cabre E, Domenech E: **Impact of environmental and dietary factors on the course of inflammatory bowel disease**. *World J Gastroenterol* 2012, **18**(29):3814-3822.
73. Faria AM, Gomes-Santos AC, Goncalves JL, Moreira TG, Medeiros SR, Dourado LP, Cara DC: **Food components and the immune system: from tonic agents to allergens**. *Frontiers in immunology* 2013, **4**:102.
74. Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, Knight R: **Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns**. *Proc Natl Acad Sci U S A* 2010, **107**(26):11971-11975.
75. Biasucci G, Benenati B, Morelli L, Bessi E, Boehm G: **Cesarean delivery may affect the early biodiversity of intestinal bacteria**. *J Nutr* 2008, **138**(9):1796s-1800s.
76. Hansen CHF, Andersen LSF, Krych L, Metzdorff SB, Hasselby JP, Skov S, Nielsen DS, Buschard K, Hansen LH, Hansen AK: **Mode of Delivery Shapes Gut Colonization Pattern and Modulates Regulatory Immunity in Mice**. *Journal of immunology* 2014, **193**(4):1213-1222.
77. Bager P, Simonsen J, Nielsen NM, Frisch M: **Cesarean section and offspring's risk of inflammatory bowel disease: A national cohort study**. *Inflammatory bowel diseases* 2012, **18**(5):857-862.
78. Ponsonby AL, Catto-Smith AG, Pezic A, Dupuis S, Halliday J, Cameron D, Morley R, Carlin J, Dwyer T: **Association Between Early-life Factors and Risk of Child-onset Crohn's Disease Among Victorian Children Born 1983-1998: A Birth Cohort Study**. *Inflammatory bowel diseases* 2009, **15**(6):858-866.
79. Bruce A, Black M, Bhattacharya S: **Mode of Delivery and Risk of Inflammatory Bowel Disease in the Offspring: Systematic Review and Meta-analysis of Observational Studies**. *Inflammatory bowel diseases* 2014, **20**(7):1217-1226.
80. Schneider V, Suissa S: **Antibiotic use and the development of Crohn's disease: methodological issues**. *Gut* 2004, **53**(10):1544-1544.
81. Card T, Logan RFA, Rodrigues LC, Wheeler JG: **Antibiotic use and the development of Crohn's disease**. *Gut* 2004, **53**(2):246-250.
82. Hildebrand H, Malmborg P, Askling J, Ekbom A, Montgomery SM: **Early-life exposures associated with antibiotic use and risk of subsequent Crohn's disease**. *Scand J Gastroentero* 2008, **43**(8):961-966.
83. Shaw SY, Blanchard JF, Bernstein CN: **Association Between the Use of Antibiotics in the First Year of Life and Pediatric Inflammatory Bowel Disease**. *American Journal of Gastroenterology* 2010, **105**(12):2687-2692.
84. Hviid A, Svanstrom H, Frisch M: **Antibiotic use and inflammatory bowel diseases in childhood**. *Gut* 2011, **60**(1):49-54.

85. Ng SC, Tang W, Leong RW, Chen M, Ko Y, Studd C, Niewiadomski O, Bell S, Kamm MA, de Silva HJ *et al*: **Environmental risk factors in inflammatory bowel disease: a population-based case-control study in Asia-Pacific**. *Gut* 2014.
86. Dethlefsen L, Relman DA: **Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation**. *Proc Natl Acad Sci U S A* 2011, **108 Suppl 1**:4554-4561.
87. Jakobsson HE, Jernberg C, Andersson AF, Sjolund-Karlsson M, Jansson JK, Engstrand L: **Short-Term Antibiotic Treatment Has Differing Long-Term Impacts on the Human Throat and Gut Microbiome**. *Plos One* 2010, **5**(3).
88. Dethlefsen L, Huse S, Sogin ML, Relman DA: **The Pervasive Effects of an Antibiotic on the Human Gut Microbiota, as Revealed by Deep 16S rRNA Sequencing**. *Plos Biology* 2008, **6**(11):2383-2400.
89. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T *et al*: **A human gut microbial gene catalogue established by metagenomic sequencing**. *Nature* 2010, **464**(7285):59-65.
90. Minot S, Bryson A, Chehoud C, Wu GD, Lewis JD, Bushman FD: **Rapid evolution of the human gut virome**. *Proc Natl Acad Sci U S A* 2013, **110**(30):12450-12455.
91. Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE, Relman DA: **Diversity of the human intestinal microbial flora**. *Science* 2005, **308**(5728):1635-1638.
92. Eckburg PB, Relman DA: **The role of microbes in Crohn's disease**. *Clin Infect Dis* 2007, **44**(2):256-262.
93. Sartor RB: **Microbial influences in inflammatory bowel diseases**. *Gastroenterology* 2008, **134**(2):577-594.
94. Frank DN, St Amand AL, Feldman RA, Boedeker EC, Harpaz N, Pace NR: **Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases**. *Proc Natl Acad Sci U S A* 2007, **104**(34):13780-13785.
95. Martinez I, Muller CE, Walter J: **Long-term temporal analysis of the human fecal microbiota revealed a stable core of dominant bacterial species**. *Plos One* 2013, **8**(7):e69621.
96. Carmody RN, Gerber GK, Luevano JM, Jr., Gatti DM, Somes L, Svenson KL, Turnbaugh PJ: **Diet dominates host genotype in shaping the murine gut microbiota**. *Cell host & microbe* 2015, **17**(1):72-84.
97. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R: **Diversity, stability and resilience of the human gut microbiota**. *Nature* 2012, **489**(7415):220-230.
98. Qin J, Li Y, Cai Z, Li S, Zhu J, Zhang F, Liang S, Zhang W, Guan Y, Shen D *et al*: **A metagenome-wide association study of gut microbiota in type 2 diabetes**. *Nature* 2012, **490**(7418):55-60.
99. Ray K: **Gut microbiota: Adding weight to the microbiota's role in obesity--exposure to antibiotics early in life can lead to increased adiposity**. *Nat Rev Gastroenterol Hepatol* 2012, **9**(11):615.
100. Claesson MJ, Cusack S, O'Sullivan O, Greene-Diniz R, de Weerd H, Flannery E, Marchesi JR, Falush D, Dinan T, Fitzgerald G *et al*: **Composition, variability, and**

temporal stability of the intestinal microbiota of the elderly. *Proc Natl Acad Sci U S A* 2011, **108 Suppl 1**:4586-4591.
101. Palmer C, Bik EM, DiGiulio DB, Relman DA, Brown PO: **Development of the human infant intestinal microbiota**. *PLoS Biol* 2007, **5**(7):e177.
102. Hopkins MJ, Sharp R, Macfarlane GT: **Age and disease related changes in intestinal bacterial populations assessed by cell culture, 16S rRNA abundance, and community cellular fatty acid profiles**. *Gut* 2001, **48**(2):198-205.
103. Mariat D, Firmesse O, Levenez F, Guimaraes V, Sokol H, Dore J, Corthier G, Furet JP: **The Firmicutes/Bacteroidetes ratio of the human microbiota changes with age**. *BMC Microbiol* 2009, **9**:123.
104. Enck P, Zimmermann K, Rusch K, Schwiertz A, Klosterhalfen S, Frick JS: **The effects of ageing on the colonic bacterial microflora in adults**. *Z Gastroenterol* 2009, **47**(7):653-658.
105. Woodmansey EJ: **Intestinal bacteria and ageing**. *J Appl Microbiol* 2007, **102**(5):1178-1186.
106. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, Magris M, Hidalgo G, Baldassano RN, Anokhin AP *et al*: **Human gut microbiome viewed across age and geography**. *Nature* 2012, **486**(7402):222-227.
107. Lozupone CA, Stombaugh J, Gonzalez A, Ackermann G, Wendel D, Vazquez-Baeza Y, Jansson JK, Gordon JI, Knight R: **Meta-analyses of studies of the human microbiota**. *Genome research* 2013, **23**(10):1704-1714.
108. Tremaroli V, Bäckhed F: **Functional interactions between the gut microbiota and host metabolism**. *Nature* 2012, **489**:242-249.
109. Cummings JH, Macfarlane GT: **The control and consequences of bacterial fermentation in the human colon**. *J Appl Bacteriol* 1991, **70**(6):443-459.
110. Martens EC, Lowe EC, Chiang H, Pudlo NA, Wu M, McNulty NP, Abbott DW, Henrissat B, Gilbert HJ, Bolam DN *et al*: **Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts**. *PLoS Biol* 2011, **9**(12):e1001221.
111. Louis P, Young P, Holtrop G, Flint HJ: **Diversity of human colonic butyrate-producing bacteria revealed by analysis of the butyryl-CoA:acetate CoA-transferase gene**. *Environ Microbiol* 2010, **12**(2):304-314.
112. Bergman EN: **Energy Contributions of Volatile Fatty-Acids from the Gastrointestinal-Tract in Various Species**. *Physiol Rev* 1990, **70**(2):567-590.
113. Wolever TMS, Spadafora P, Eshuis H: **Interaction between Colonic Acetate and Propionate in Humans**. *Am J Clin Nutr* 1991, **53**(3):681-687.
114. Donohoe DR, Garge N, Zhang X, Sun W, O'Connell TM, Bunger MK, Bultman SJ: **The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon**. *Cell Metab* 2011, **13**(5):517-526.
115. Barcelo A, Claustre J, Moro F, Chayvialle JA, Cuber JC, Plaisancie P: **Mucin secretion is modulated by luminal factors in the isolated vascularly perfused rat colon**. *Gut* 2000, **46**(2):218-224.
116. Hudcovic T, Kolinska J, Klepetar J, Stepankova R, Rezanka T, Srutkova D, Schwarzer M, Erban V, Du Z, Wells JM *et al*: **Protective effect of Clostridium tyrobutyricum in**

acute dextran sodium sulphate-induced colitis: differential regulation of tumour necrosis factor-alpha and interleukin-18 in BALB/c and severe combined immunodeficiency mice. *Clin Exp Immunol* 2012, **167**(2):356-365.

117. Bordin M, D'Atri F, Guillemot L, Citi S: **Histone deacetylase inhibitors up-regulate the expression of tight junction proteins**. *Molecular cancer research : MCR* 2004, **2**(12):692-701.
118. Petersson J, Schreiber O, Hansson GC, Gendler SJ, Velcich A, Lundberg JO, Roos S, Holm L, Phillipson M: **Importance and regulation of the colonic mucus barrier in a mouse model of colitis**. *Am J Physiol Gastrointest Liver Physiol* 2011, **300**(2):G327-333.
119. Lee YK, Mazmanian SK: **Has the Microbiota Played a Critical Role in the Evolution of the Adaptive Immune System?** *Science* 2010, **330**(6012):1768-1773.
120. Olszak T, An DD, Zeissig S, Vera MP, Richter J, Franke A, Glickman JN, Siebert R, Baron RM, Kasper DL *et al*: **Microbial Exposure During Early Life Has Persistent Effects on Natural Killer T Cell Function**. *Science* 2012, **336**(6080):489-493.
121. Strauch UG, Obermeier F, Grunwald N, Gurster S, Dunger N, Schultz M, Griese DP, Mahler M, Scholmerich J, Rath HC: **Influence of intestinal bacteria on induction of regulatory T cells: lessons from a transfer model of colitis**. *Gut* 2005, **54**(11):1546-1552.
122. Chung H, Pamp SJ, Hill JA, Surana NK, Edelman SM, Troy EB, Reading NC, Villablanca EJ, Wang S, Mora JR *et al*: **Gut immune maturation depends on colonization with a host-specific microbiota**. *Cell* 2012, **149**(7):1578-1593.
123. Saulnier DM, Riehle K, Mistretta TA, Diaz MA, Mandal D, Raza S, Weidler EM, Qin X, Coarfa C, Milosavljevic A *et al*: **Gastrointestinal microbiome signatures of pediatric patients with irritable bowel syndrome**. *Gastroenterology* 2011, **141**(5):1782-1791.
124. Greenblum S, Turnbaugh PJ, Borenstein E: **Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease**. *Proc Natl Acad Sci U S A* 2012, **109**(2):594-599.
125. Tilg H, Moschen AR: **Microbiota and diabetes: an evolving relationship**. *Gut* 2014.
126. Garrett WS, Lord GM, Punit S, Lugo-Villarino G, Mazmanian SK, Ito S, Glickman JN, Glimcher LH: **Communicable ulcerative colitis induced by T-bet deficiency in the innate immune system**. *Cell* 2007, **131**(1):33-45.
127. Baumgart M, Dogan B, Rishniw M, Weitzman G, Bosworth B, Yantiss R, Orsi RH, Wiedmann M, McDonough P, Kim SG *et al*: **Culture independent analysis of ileal mucosa reveals a selective increase in invasive Escherichia coli of novel phylogeny relative to depletion of Clostridiales in Crohn's disease involving the ileum**. *The ISME journal* 2007, **1**(5):403-418.
128. Ott SJ, Musfeldt M, Wenderoth DF, Hampe J, Brant O, Folsch UR, Timmis KN, Schreiber S: **Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease**. *Gut* 2004, **53**(5):685-693.
129. Swidsinski A, Weber J, Loening-Baucke V, Hale LP, Lochs H: **Spatial organization and composition of the mucosal flora in patients with inflammatory bowel disease**. *J Clin Microbiol* 2005, **43**(7):3380-3389.

130. Sokol H, Seksik P, Furet JP, Firmesse O, Nion-Larmurier I, Beaugerie L, Cosnes J, Corthier G, Marteau P, Dore J: **Low counts of Faecalibacterium prausnitzii in colitis microbiota**. *Inflammatory bowel diseases* 2009, **15**(8):1183-1189.
131. Sokol H, Pigneur B, Watterlot L, Lakhdari O, Bermudez-Humaran LG, Gratadoux JJ, Blugeon S, Bridonneau C, Furet JP, Corthier G *et al*: **Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients**. *Proc Natl Acad Sci U S A* 2008, **105**(43):16731-16736.
132. Willing BP, Dicksved J, Halfvarson J, Andersson AF, Lucio M, Zheng Z, Jarnerot G, Tysk C, Jansson JK, Engstrand L: **A Pyrosequencing Study in Twins Shows That Gastrointestinal Microbial Profiles Vary With Inflammatory Bowel Disease Phenotypes**. *Gastroenterology* 2010, **139**(6):1844-U1105.
133. Lepage P, Hasler R, Spehlmann ME, Rehman A, Zvirbliene A, Begun A, Ott S, Kupcinskas L, Dore J, Raedler A *et al*: **Twin Study Indicates Loss of Interaction Between Microbiota and Mucosa of Patients With Ulcerative Colitis**. *Gastroenterology* 2011, **141**(1):227-236.
134. Conte MP, Schippa S, Zamboni I, Penta M, Chiarini F, Seganti L, Osborn J, Falconieri P, Borrelli O, Cucchiara S: **Gut-associated bacterial microbiota in paediatric patients with inflammatory bowel disease**. *Gut* 2006, **55**(12):1760-1767.
135. Swidsinski A, Ladhoff A, Pernthaler A, Swidsinski S, Loening-Baucke V, Ortner M, Weber J, Hoffmann U, Schreiber S, Dietel M *et al*: **Mucosal flora in inflammatory bowel disease**. *Gastroenterology* 2002, **122**(1):44-54.
136. Schwiertz A, Jacobi M, Frick JS, Richter M, Rusch K, Kohler H: **Microbiota in pediatric inflammatory bowel disease**. *J Pediatr* 2010, **157**(2):240-244 e241.
137. Cucchiara S, Iebba V, Conte MP, Schippa S: **The microbiota in inflammatory bowel disease in different age groups**. *Dig Dis* 2009, **27**(3):252-258.
138. Papa E, Docktor M, Smillie C, Weber S, Preheim SP, Gevers D, Giannoukos G, Ciulla D, Tabbaa D, Ingram J *et al*: **Non-Invasive Mapping of the Gastrointestinal Microbiota Identifies Children with Inflammatory Bowel Disease**. *Plos One* 2012, **7**(6).
139. Gevers D, Kugathasan S, Denson LA, Vazquez-Baeza Y, Van Treuren W, Ren B, Schwager E, Knights D, Song SJ, Yassour M *et al*: **The treatment-naive microbiome in new-onset Crohn's disease**. *Cell host & microbe* 2014, **15**(3):382-392.
140. Hansen R, Russell RK, Reiff C, Louis P, McIntosh F, Berry SH, Mukhopadhya I, Bisset WM, Barclay AR, Bishop J *et al*: **Microbiota of De-Novo Pediatric IBD: Increased Faecalibacterium Prausnitzii and Reduced Bacterial Diversity in Crohn's But Not in Ulcerative Colitis**. *American Journal of Gastroenterology* 2012, **107**(12):1913-1922.
141. Schwiertz A, Jacobi M, Frick JS, Richter M, Rusch K, Kohler H: **Microbiota in Pediatric Inflammatory Bowel Disease**. *J Pediatr* 2010.
142. Takaishi H, Matsuki T, Nakazawa A, Takada T, Kado S, Asahara T, Kamada N, Sakuraba A, Yajima T, Higuchi H *et al*: **Imbalance in intestinal microflora constitution could be involved in the pathogenesis of inflammatory bowel disease**. *Int J Med Microbiol* 2008, **298**(5-6):463-472.

143. Li M, Wang B, Zhang M, Rantalainen M, Wang S, Zhou H, Zhang Y, Shen J, Pang X, Wei H *et al*: **Symbiotic gut microbes modulate human metabolic phenotypes**. *Proc Natl Acad Sci U S A* 2008, **105**(6):2117-2122.
144. Zhang H, DiBaise JK, Zuccolo A, Kudrna D, Braidotti M, Yu Y, Parameswaran P, Crowell MD, Wing R, Rittmann BE *et al*: **Human gut microbiota in obesity and after gastric bypass**. *Proc Natl Acad Sci U S A* 2009, **106**(7):2365-2370.
145. Tap J, Mondot S, Levenez F, Pelletier E, Caron C, Furet JP, Ugarte E, Munoz-Tamayo R, Paslier DL, Nalin R *et al*: **Towards the human intestinal microbiota phylogenetic core**. *Environ Microbiol* 2009, **11**(10):2574-2584.
146. Turnbaugh PJ, Ridaura VK, Faith JJ, Rey FE, Knight R, Gordon JI: **The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice**. *Sci Transl Med* 2009, **1**(6):6ra14.
147. Hansen R, Russell RK, Reiff C, Louis P, McIntosh F, Berry SH, Mukhopadhya I, Bisset WM, Barclay AR, Bishop J *et al*: **Microbiota of de-novo pediatric IBD: increased Faecalibacterium prausnitzii and reduced bacterial diversity in Crohn's but not in ulcerative colitis**. *Am J Gastroenterol* 2012, **107**(12):1913-1922.
148. Hong PY, Croix JA, Greenberg E, Gaskins HR, Mackie RI: **Pyrosequencing-based analysis of the mucosal microbiota in healthy individuals reveals ubiquitous bacterial groups and micro-heterogeneity**. *Plos One* 2011, **6**(9):e25042.
149. van den Bogert B, de Vos WM, Zoetendal EG, Kleerebezem M: **Microarray analysis and barcoded pyrosequencing provide consistent microbial profiles depending on the source of human intestinal samples**. *Appl Environ Microbiol* 2011, **77**(6):2071-2080.
150. Hansson GC: **Role of mucus layers in gut infection and inflammation**. *Current Opinion in Microbiology* 2012, **15**(1):57-62.
151. Harrell L, Wang Y, Antonopoulos D, Young V, Lichtenstein L, Huang Y, Hanauer S, Chang E: **Standard colonic lavage alters the natural state of mucosal-associated microbiota in the human colon**. *Plos One* 2012, **7**(2):e32545.
152. Johansson ME, Ambort D, Pelaseyed T, Schutte A, Gustafsson JK, Ermund A, Subramani DB, Holmen-Larsson JM, Thomsson KA, Bergstrom JH *et al*: **Composition and functional role of the mucus layers in the intestine**. *Cellular and molecular life sciences : CMLS* 2011, **68**(22):3635-3641.
153. Huse SM, Young VB, Morrison HG, Antonopoulos DA, Kwon J, Dalal S, Arrieta R, Hubert NA, Shen L, Vineis JH *et al*: **Comparison of brush and biopsy sampling methods of the ileal pouch for assessment of mucosa-associated microbiota of human subjects**. *Microbiome* 2014, **2**(1):5.
154. Clement BG, Kehl LE, DeBord KL, Kitts CL: **Terminal restriction fragment patterns (TRFPs), a rapid, PCR-based method for the comparison of complex bacterial communities**. *J Microbiol Meth* 1998, **31**(3):135-142.
155. Muyzer G, Smalla K: **Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology**. *Antonie van Leeuwenhoek* 1998, **73**(1):127-141.

156. Armougom F, Raoult D: **Use of pyrosequencing and DNA barcodes to monitor variations in *Firmicutes* and *Bacteroidetes* communities in the gut microbiota of obese humans**. *BMC Genomics* 2008, **9**:576.
157. Claesson MJ, O'Sullivan O, Wang Q, Nikkila J, Marchesi JR, Smidt H, de Vos WM, Ross RP, O'Toole PW: **Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine**. *Plos One* 2009, **4**(8):e6669.
158. Larsen N, Vogensen FK, van den Berg FW, Nielsen DS, Andreasen AS, Pedersen BK, Al-Soud WA, Sorensen SJ, Hansen LH, Jakobsen M: **Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults**. *Plos One* 2010, **5**(2):e9085.
159. Harrington CR, Lucchini S, Ridgway KP, Wegmann U, Eaton TJ, Hinton JCD, Gasson MJ, Narbad A: **A short-oligonucleotide microarray that allows improved detection of gastrointestinal tract microbial communities**. *Bmc Microbiology* 2008, **8**.
160. Palmer C, Bik EM, Eisen MB, Eckburg PB, Sana TR, Wolber PK, Relman DA, Brown PO: **Rapid quantitative profiling of complex microbial populations**. *Nucleic Acids Res* 2006, **34**(1).
161. Wang RF, Beggs ML, Erickson BD, Cerniglia CE: **DNA microarray analysis of predominant human intestinal bacteria in fecal samples**. *Mol Cell Probes* 2004, **18**(4):223-234.
162. Wang RF, Beggs ML, Robertson LH, Cerniglia CE: **Design and evaluation of oligonucleotide-microarray method for the detection of human intestinal bacteria in fecal samples**. *FEMS Microbiol Lett* 2002, **213**(2):175-182.
163. Wilson KH, Blitchington RB: **Human colonic biota studied by ribosomal DNA sequence analysis**. *Appl Environ Microbiol* 1996, **62**(7):2273-2278.
164. Van de Peer Y, Chapelle S, De Wachter R: **A quantitative map of nucleotide substitution rates in bacterial rRNA**. *Nucleic Acids Res* 1996, **24**(17):3381-3391.
165. Chakravorty S, Helb D, Burday M, Connell N, Alland D: **A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria**. *J Microbiol Methods* 2007, **69**(2):330-339.
166. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R: **Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample**. *P Natl Acad Sci USA* 2011, **108**:4516-4522.
167. Arthur JC, Perez-Chanona E, Muhlbauer M, Tomkovich S, Uronis JM, Fan TJ, Campbell BJ, Abujamel T, Dogan B, Rogers AB *et al*: **Intestinal inflammation targets cancer-inducing activity of the microbiota**. *Science* 2012, **338**(6103):120-123.
168. Joossens M, Huys G, Cnockaert M, De Preter V, Verbeke K, Rutgeerts P, Vandamme P, Vermeire S: **Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives**. *Gut* 2011, **60**(5):631-637.
169. Tong M, Li X, Wegener Parfrey L, Roth B, Ippoliti A, Wei B, Borneman J, McGovern DP, Frank DN, Li E *et al*: **A modular organization of the human intestinal mucosal microbiota and its association with inflammatory bowel disease**. *Plos One* 2013, **8**(11):e80702.

170. Carbonero F, Benefiel AC, Gaskins HR: **Contributions of the microbial hydrogen economy to colonic homeostasis**. *Nat Rev Gastroenterol Hepatol* 2012, **9**(9):504-518.
171. Awano N, Wada M, Mori H, Nakamori S, Takagi H: **Identification and functional analysis of Escherichia coli cysteine desulfhydrases**. *Appl Environ Microbiol* 2005, **71**(7):4149-4152.
172. Blachier F, Davila AM, Mimoun S, Benetti PH, Atanasiu C, Andriamihaja M, Benamouzig R, Bouillaud F, Tome D: **Luminal sulfide and large intestine mucosa: friend or foe?** *Amino Acids* 2010, **39**(2):335-347.
173. Loubinoux J, Bronowicki JP, Pereira IAC, Mougenel JL, Le Faou AE: **Sulfate-reducing bacteria in human feces and their association with inflammatory bowel diseases**. *Fems Microbiol Ecol* 2002, **40**(2):107-112.
174. Scanlan PD, Shanahan F, Marchesi JR: **Culture-independent analysis of desulfovibrios in the human distal colon of healthy, colorectal cancer and polypectomized individuals**. *Fems Microbiol Ecol* 2009, **69**(2):213-221.
175. Jia W, Whitehead RN, Griffiths L, Dawson C, Bai H, Waring RH, Ramsden DB, Hunter JO, Cauchi M, Bessant C *et al*: **Diversity and distribution of sulphate-reducing bacteria in human faeces from healthy subjects and patients with inflammatory bowel disease**. *FEMS Immunol Med Microbiol* 2012, **65**(1):55-68.
176. Washio J, Sato T, Koseki T, Takahashi N: **Hydrogen sulfide-producing bacteria in tongue biofilm and their relationship with oral malodour**. *J Med Microbiol* 2005, **54**(9):889-895.
177. Shibuya N, Koike S, Tanaka M, Ishigami-Yuasa M, Kimura Y, Ogasawara Y, Fukui K, Nagahara N, Kimura H: **A novel pathway for the production of hydrogen sulfide from D-cysteine in mammalian cells**. *Nat Commun* 2013, **4**:1366.
178. Macfarlane GT, Gibson GR, Cummings JH: **Comparison of fermentation reactions in different regions of the human colon**. *J Appl Bacteriol* 1992, **72**(1):57-64.
179. Florin TH: **Hydrogen sulphide and total acid-volatile sulphide in faeces, determined with a direct spectrophotometric method**. *Clin Chim Acta* 1991, **196**(2-3):127-134.
180. Magee EA, Richardson CJ, Hughes R, Cummings JH: **Contribution of dietary protein to sulfide production in the large intestine: an in vitro and a controlled feeding study in humans**. *Am J Clin Nutr* 2000, **72**(6):1488-1494.
181. Wallace JL, Blackler RW, Chan MV, Da Silva GJ, Elsheikh W, Flannigan KL, Gamaniek I, Manko A, Wang L, Motta JP *et al*: **Anti-Inflammatory and Cytoprotective Actions of Hydrogen Sulfide: Translation to Therapeutics**. *Antioxid Redox Signal* 2014.
182. Rose P, Moore PK, Ming SH, Nam OC, Armstrong JS, Whiteman M: **Hydrogen sulfide protects colon cancer cells from chemopreventative agent beta-phenylethyl isothiocyanate induced apoptosis**. *World J Gastroenterol* 2005, **11**(26):3990-3997.
183. Zanardo RCO, Brancaleone V, Distrutti E, Fiorucci S, Cirino G, Wallace JL: **Hydrogen sulfide is an endogenous modulator of leukocyte-mediated inflammation**. *Faseb J* 2006, **20**(12):2118-+.
184. Wallace JL, Vong L, McKnight W, Dicay M, Martin GR: **Endogenous and exogenous hydrogen sulfide promotes resolution of colitis in rats**. *Gastroenterology* 2009, **137**(2):569-578, 578 e561.

185. Mariggio MA, Minunno V, Riccardi S, Santacroce R, De Rinaldis P, Fumarulo R: **Sulfide enhancement of PMN apoptosis**. *Immunopharmacol Immunotoxicol* 1998, **20**(3):399-408.
186. Rinaldi L, Gobbi G, Pambianco M, Micheloni C, Mirandola P, Vitale M: **Hydrogen sulfide prevents apoptosis of human PMN via inhibition of p38 and caspase 3**. *Lab Invest* 2006, **86**(4):391-397.
187. Wallace JL, Dicay M, McKnight W, Martin GR: **Hydrogen sulfide enhances ulcer healing in rats**. *Faseb J* 2007, **21**(14):4070-4076.
188. Wu YC, Wang XJ, Yu L, Chan FK, Cheng AS, Yu J, Sung JJ, Wu WK, Cho CH: **Hydrogen sulfide lowers proliferation and induces protective autophagy in colon epithelial cells**. *Plos One* 2012, **7**(5):e37572.
189. Attene-Ramos MS, Nava GM, Muellner MG, Wagner ED, Plewa MJ, Gaskins HR: **DNA Damage and Toxicogenomic Analyses of Hydrogen Sulfide in Human Intestinal Epithelial FHs 74 Int Cells**. *Environ Mol Mutagen* 2010, **51**(4):304-314.
190. Leschelle X, Goubern M, Andriamihaja M, Blottiere HM, Couplan E, Gonzalez-Barroso MD, Petit C, Pagniez A, Chaumontet C, Mignotte B *et al*: **Adaptative metabolic response of human colonic epithelial cells to the adverse effects of the luminal compound sulfide**. *Biochim Biophys Acta* 2005, **1725**(2):201-212.
191. Roediger WE, Duncan A, Kapaniris O, Millard S: **Sulphide impairment of substrate oxidation in rat colonocytes: a biochemical basis for ulcerative colitis?** *Clin Sci (Lond)* 1993, **85**(5):623-627.
192. Moore JW, Babidge W, Millard S, Roediger WE: **Effect of sulphide on short chain acyl-CoA metabolism in rat colonocytes**. *Gut* 1997, **41**(1):77-81.
193. Christl SU, Eisner HD, Dusel G, Kasper H, Scheppach W: **Antagonistic effects of sulfide and butyrate on proliferation of colonic mucosa - A potential role for these agents in the pathogenesis of ulcerative colitis**. *Digest Dis Sci* 1996, **41**(12):2477-2481.
194. Vermeiren J, Van de Wiele T, Van Nieuwenhuyse G, Boeckx P, Verstraete W, Boon N: **Sulfide- and nitrite-dependent nitric oxide production in the intestinal tract**. *Microb Biotechnol* 2012, **5**(3):379-387.
195. Furne J, Springfield J, Koenig T, DeMaster E, Levitt MD: **Oxidation of hydrogen sulfide and methanethiol to thiosulfate by rat tissues: a specialized function of the colonic mucosa**. *Biochem Pharmacol* 2001, **62**(2):255-259.
196. Mimoun S, Andriamihaja M, Chaumontet C, Atanasiu C, Benamouzig R, Blouin JM, Tome D, Bouillaud F, Blachier F: **Detoxification of H2S by Differentiated Colonic Epithelial Cells: Implication of the Sulfide Oxidizing Unit and of the Cell Respiratory Capacity**. *Antioxid Redox Sign* 2012, **17**(1):1-10.
197. Turnbaugh PJ: **Microbiology: fat, bile and gut microbes**. *Nature* 2012, **487**(7405):47-48.
198. Levine J, Ellis CJ, Furne JK, Springfield J, Levitt MD: **Fecal hydrogen sulfide production in ulcerative colitis**. *The American journal of gastroenterology* 1998, **93**(1):83-87.

199. Edmond LM, Hopkins MJ, Magee EA, Cummings JH: **The effect of 5-aminosalicylic acid-containing drugs on sulfide production by sulfate-reducing and amino acid-fermenting bacteria**. *Inflammatory bowel diseases* 2003, **9**(1):10-17.
200. Duffy M OML, Coffey JC, Collins JK, Shanahan F, Redmond HP, Kirwan WO: **Sulfate-reducing bacteria colonize pouches formed for ulcerative colitis but not for familial adenomatous polyposis**. *Dis Colon Rectum* 2002, **45**(3):384-388.
201. Moore J BW, Millard S, Roediger W.: **Colonic luminal hydrogen sulfide is not elevated in ulcerative colitis.** *Digest Dis Sci* 1998, **43**(1):162-165.
202. Pitcher MC, Beatty ER, Cummings JH: **The contribution of sulphate reducing bacteria and 5-aminosalicylic acid to faecal sulphide in patients with ulcerative colitis**. *Gut* 2000, **46**(1):64-72.
203. Zinkevich VV, Beech IB: **Screening of sulfate-reducing bacteria in colonoscopy samples from healthy and colitic human gut mucosa**. *Fems Microbiol Ecol* 2000, **34**(2):147-155.
204. Ng SC, Bernstein CN, Vatn MH, Lakatos PL, Loftus EV, Jr., Tysk C, O'Morain C, Moum B, Colombel JF: **Geographical variability and environmental risk factors in inflammatory bowel disease**. *Gut* 2013, **62**(4):630-649.
205. Tsai HH, Dwarakanath AD, Hart CA, Milton JD, Rhodes JM: **Increased faecal mucin sulphatase activity in ulcerative colitis: a potential target for treatment**. *Gut* 1995, **36**(4):570-576.
206. Pitcher MC, Beatty ER, Harris RM, Waring RH, Cummings JH: **Sulfur metabolism in ulcerative colitis: investigation of detoxification enzymes in peripheral blood**. *Dig Dis Sci* 1998, **43**(9):2080-2085.
207. Picton R, Eggo MC, Langman MJ, Singh S: **Impaired detoxication of hydrogen sulfide in ulcerative colitis?** *Dig Dis Sci* 2007, **52**(2):373-378.
208. Taniguchi E, Matsunami M, Kimura T, Yonezawa D, Ishiki T, Sekiguchi F, Nishikawa H, Maeda Y, Ishikura H, Kawabata A: **Rhodanese, but not cystathionine-gamma-lyase, is associated with dextran sulfate sodium-evoked colitis in mice: A sign of impaired colonic sulfide detoxification?** *Toxicology* 2009, **264**(1-2):96-103.
209. De Preter V, Arijs I, Windey K, Vanhove W, Vermeire S, Schuit F, Rutgeerts P, Verbeke K: **Decreased mucosal sulfide detoxification is related to an impaired butyrate oxidation in ulcerative colitis**. *Inflammatory bowel diseases* 2012, **18**(12):2371-2380.
210. Louis P, Flint HJ: **Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine**. *FEMS Microbiol Lett* 2009, **294**(1):1-8.
211. Duncan SH, Barcenilla A, Stewart CS, Pryde SE, Flint HJ: **Acetate utilization and butyryl coenzyme A (CoA): acetate-CoA transferase in butyrate-producing bacteria from the human large intestine**. *Appl Environ Microb* 2002, **68**(10):5186-5190.
212. Walker AW, Duncan SH, Leitch ECM, Child MW, Flint HJ: **pH and peptide supply can radically alter bacterial populations and short-chain fatty acid ratios within microbial communities from the human colon**. *Appl Environ Microb* 2005, **71**(7):3692-3700.

213. Fallingborg J: **Intraluminal pH of the human gastrointestinal tract**. *Danish medical bulletin* 1999, **46**(3):183-196.
214. De Preter V, Arijs I, Windey K, Vanhove W, Vermeire S, Schuit F, Rutgeerts P, Verbeke K: **Impaired butyrate oxidation in ulcerative colitis is due to decreased butyrate uptake and a defect in the oxidation pathway**. *Inflammatory bowel diseases* 2012, **18**(6):1127-1136.
215. De Preter V, Geboes KP, Bulteel V, Vandermeulen G, Suenaert P, Rutgeerts P, Verbeke K: **Kinetics of butyrate metabolism in the normal colon and in ulcerative colitis: the effects of substrate concentration and carnitine on the beta-oxidation pathway**. *Aliment Pharm Ther* 2011, **34**(5):526-532.
216. Kovarik JJ, Tillinger W, Hofer J, Holzl MA, Heinzl H, Saemann MD, Zlabinger GJ: **Impaired anti-inflammatory efficacy of n-butyrate in patients with IBD**. *Eur J Clin Invest* 2011, **41**(3):291-298.
217. Roediger WE: **The colonic epithelium in ulcerative colitis: an energy-deficiency disease?** *Lancet* 1980, **2**(8197):712-715.
218. Roediger WEW, Heyworth M, Willoughby P, Piris J, Moore A, Truelove SC: **Luminal Ions and Short Chain Fatty-Acids as Markers of Functional-Activity of the Mucosa in Ulcerative-Colitis**. *J Clin Pathol* 1982, **35**(3):323-326.
219. Thibault R, De Coppet P, Daly K, Bourreille A, Cuff M, Bonnet C, Mosnier JF, Galmiche JP, Shirazi-Beechey S, Segain JP: **Down-regulation of the monocarboxylate transporter 1 is involved in butyrate deficiency during intestinal inflammation**. *Gastroenterology* 2007, **133**(6):1916-1927.
220. Butzner JD, Parmar R, Bell CJ, Dalal V: **Butyrate enema therapy stimulates mucosal repair in experimental colitis in the rat**. *Gut* 1996, **38**(4):568-573.
221. Vieira ELM, Leonel AJ, Sad AP, Beltrao NRM, Costa TF, Ferreira TMR, Gomes-Santos AC, Faria AMC, Peluzio MCG, Cara DC *et al*: **Oral administration of sodium butyrate attenuates inflammation and mucosal lesion in experimental acute ulcerative colitis**. *J Nutr Biochem* 2012, **23**(5):430-436.
222. Segain JP, de la Bletiere DR, Bourreille A, Leray V, Gervois N, Rosales C, Ferrier L, Bonnet C, Blottiere HM, Galmiche JP: **Butyrate inhibits inflammatory responses through NF kappa B inhibition: implications for Crohn's disease**. *Gut* 2000, **47**(3):397-403.
223. Luhrs H, Gerke T, Muller JG, Melcher R, Schauber J, Boxberge F, Scheppach W, Menzel T: **Butyrate inhibits NF-kappaB activation in lamina propria macrophages of patients with ulcerative colitis**. *Scand J Gastroenterol* 2002, **37**(4):458-466.
224. Zimmerman MA, Singh N, Martin PM, Thangaraju M, Ganapathy V, Waller JL, Shi HD, Robertson KD, Munn DH, Liu KB: **Butyrate suppresses colonic inflammation through HDAC1-dependent Fas upregulation and Fas-mediated apoptosis of T cells**. *Am J Physiol-Gastr L* 2012, **302**(12):G1405-G1415.
225. Russo I, Luciani A, De Cicco P, Troncone E, Ciacci C: **Butyrate Attenuates Lipopolysaccharide-Induced Inflammation in Intestinal Cells and Crohn's Mucosa through Modulation of Antioxidant Defense Machinery**. *Plos One* 2012, **7**(3).

226. Sauer J, Richter KK, Pool-Zobel BL: **Physiological concentrations of butyrate favorably modulate genes of oxidative and metabolic stress in primary human colon cells**. *J Nutr Biochem* 2007, **18**(11):736-745.
227. Manichanh C, Rigottier-Gois L, Bonnaud E, Gloux K, Pelletier E, Frangeul L, Nalin R, Jarrin C, Chardon P, Marteau P *et al*: **Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach**. *Gut* 2006, **55**(2):205-211.
228. Sokol H, Pigneur B, Watterlot L, Lakhdari O, Bermudez-Humaran LG, Gratadoux JJ, Blugeon S, Bridonneau C, Furet JP, Corthier G *et al*: **Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients**. *Proceedings of the National Academy of Sciences of the United States of America* 2008, **105**(43):16731-16736.
229. Kumari R, Ahuja V, Paul J: **Fluctuations in butyrate-producing bacteria in ulcerative colitis patients of North India**. *World J Gastroenterol* 2013, **19**(22):3404-3414.
230. North American Society for Pediatric Gastroenterology H, Nutrition, Colitis Foundation of A, Bousvaros A, Antonioli DA, Colletti RB, Dubinsky MC, Glickman JN, Gold BD, Griffiths AM *et al*: **Differentiating ulcerative colitis from Crohn disease in children and young adults: report of a working group of the North American Society for Pediatric Gastroenterology, Hepatology, and Nutrition and the Crohn's and Colitis Foundation of America**. *J Pediatr Gastroenterol Nutr* 2007, **44**(5):653-674.
231. Daperno M, D'Haens G, Van Assche G, Baert F, Bulois P, Maunoury V, Sostegni R, Rocca R, Pera A, Gevers A *et al*: **Development and validation of a new, simplified endoscopic activity score for Crohn's disease: the SES-CD**. *Gastrointestinal endoscopy* 2004, **60**(4):505-512.
232. Levine A, Griffiths A, Markowitz J, Wilson DC, Turner D, Russell RK, Fell J, Ruemmele FM, Walters T, Sherlock M *et al*: **Pediatric modification of the Montreal classification for inflammatory bowel disease: the Paris classification**. *Inflammatory bowel diseases* 2011, **17**(6):1314-1321.
233. Hyams J, Markowitz J, Otley A, Rosh J, Mack D, Bousvaros A, Kugathasan S, Pfefferkorn M, Tolia V, Evans J *et al*: **Evaluation of the pediatric crohn disease activity index: a prospective multicenter experience**. *J Pediatr Gastroenterol Nutr* 2005, **41**(4):416-421.
234. Turner D, Otley AR, Mack D, Hyams J, de Bruijne J, Uusoue K, Walters TD, Zachos M, Mamula P, Beaton DE *et al*: **Development, validation, and evaluation of a pediatric ulcerative colitis activity index: a prospective multicenter study**. *Gastroenterology* 2007, **133**(2):423-432.
235. Jimenez-Rivera C, Haas D, Boland M, Barkey JL, Mack DR: **Comparison of two common outpatient preparations for colonoscopy in children and youth**. *Gastroenterology research and practice* 2009, **2009**:518932.
236. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI *et al*: **QIIME allows analysis of high-throughput community sequencing data**. *Nature methods* 2010, **7**(5):335-336.
237. Edgar RC: **Search and clustering orders of magnitude faster than BLAST**. *Bioinformatics* 2010, **26**(19):2460-2461.

238. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool**. *Journal of molecular biology* 1990, **215**(3):403-410.
239. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C: **Metagenomic biomarker discovery and explanation**. *Genome Biol* 2011, **12**(6):R60.
240. Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ, Kulam-Syed-Mohideen AS, McGarrell DM, Marsh T, Garrity GM *et al*: **The Ribosomal Database Project: improved alignments and new tools for rRNA analysis**. *Nucleic Acids Res* 2009, **37**(Database issue):D141-145.
241. Sundquist A, Bigdeli S, Jalili R, Druzin ML, Waller S, Pullen KM, El-Sayed YY, Taslimi MM, Batzoglou S, Ronaghi M: **Bacterial flora-typing with targeted, chip-based Pyrosequencing**. *BMC Microbiol* 2007, **7**:108.
242. Magoc T, Salzberg SL: **FLASH: fast length adjustment of short reads to improve genome assemblies**. *Bioinformatics* 2011, **27**(21):2957-2963.
243. Letunic I, Bork P: **Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy**. *Nucleic Acids Res* 2011, **39**:W475-W478.
244. Langille MGI, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, Clemente JC, Burkepile DE, Thurber RLV, Knight R *et al*: **Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences**. *Nat Biotechnol* 2013, **31**(9):814-+.
245. Kostic AD, Chun E, Robertson L, Glickman JN, Gallini CA, Michaud M, Clancy TE, Chung DC, Lochhead P, Hold GL *et al*: **Fusobacterium nucleatum potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment**. *Cell host & microbe* 2013, **14**(2):207-215.
246. Boutaga K, van Winkelhoff AJ, Vandenbroucke-Grauls CM, Savelkoul PH: **Periodontal pathogens: a quantitative comparison of anaerobic culture and real-time PCR**. *FEMS Immunol Med Microbiol* 2005, **45**(2):191-199.
247. Rinttila T, Kassinen A, Malinen E, Krogius L, Palva A: **Development of an extensive set of 16S rDNA-targeted primers for quantification of pathogenic and indigenous bacteria in faecal samples by real-time PCR**. *J Appl Microbiol* 2004, **97**(6):1166-1177.
248. Bourne DG, Muirhead A, Sato Y: **Changes in sulfate-reducing bacterial populations during the onset of black band disease**. *The ISME journal* 2011, **5**(3):559-564.
249. Cox JH, Kljavin NM, Ota N, Leonard J, Roose-Girma M, Diehl L, Ouyang W, Ghilardi N: **Opposing consequences of IL-23 signaling mediated by innate and adaptive cells in chemically induced colitis in mice**. *Mucosal Immunol* 2012, **5**(1):99-109.
250. Wirtz S, Neufert C, Weigmann B, Neurath MF: **Chemically induced mouse models of intestinal inflammation**. *Nat Protoc* 2007, **2**(3):541-546.
251. Abe K, Nguyen KP, Fine SD, Mo JH, Shen C, Shenouda S, Corr M, Jung S, Lee J, Eckmann L *et al*: **Conventional dendritic cells regulate the outcome of colonic inflammation independently of T cells**. *Proc Natl Acad Sci U S A* 2007, **104**(43):17022-17027.
252. Sun X, Threadgill D, Jobin C: **Campylobacter jejuni induces colitis through activation of mammalian target of rapamycin signaling**. *Gastroenterology* 2012, **142**(1):86-95 e85.

253. Karrasch T, Kim JS, Muhlbauer M, Magness ST, Jobin C: **Gnotobiotic IL-10-/-;NF-kappa B(EGFP) mice reveal the critical role of TLR/NF-kappa B signaling in commensal bacteria-induced colitis**. *Journal of immunology* 2007, **178**(10):6522-6532.
254. Lippert E, Karrasch T, Sun X, Allard B, Herfarth HH, Threadgill D, Jobin C: **Gnotobiotic IL-10; NF-kappaB mice develop rapid and severe colitis following Campylobacter jejuni infection**. *Plos One* 2009, **4**(10):e7413.
255. Wisniewski JR, Zougman A, Nagaraj N, Mann M: **Universal sample preparation method for proteome analysis**. *Nature methods* 2009, **6**(5):359-362.
256. Stahl M, Friis LM, Nothaft H, Liu X, Li J, Szymanski CM, Stintzi A: **L-fucose utilization provides Campylobacter jejuni with a competitive advantage**. *Proc Natl Acad Sci U S A* 2011, **108**(17):7194-7199.
257. Palyada K, Sun YQ, Flint A, Butcher J, Naikare H, Stintzi A: **Characterization of the oxidative stress stimulon and PerR regulon of Campylobacter jejuni**. *BMC Genomics* 2009, **10**:481.
258. Palyada K, Threadgill D, Stintzi A: **Iron acquisition and regulation in Campylobacter jejuni**. *J Bacteriol* 2004, **186**(14):4714-4729.
259. Reid AN, Pandey R, Palyada K, Naikare H, Stintzi A: **Identification of Campylobacter jejuni genes involved in the response to acidic pH and stomach transit**. *Appl Environ Microbiol* 2008, **74**(5):1583-1597.
260. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL: **Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction**. *BMC bioinformatics* 2012, **13**:134.
261. Zhang PJ, Wei R, Wen XY, Ping L, Wang CB, Dong ZN, Deng XX, Bo W, Bin C, Tian YP: **Genes expression profiling of peripheral blood cells of patients with hepatocellular carcinoma**. *Cell biology international* 2012, **36**(9):803-809.
262. Agans R, Rigsbee L, Kenche H, Michail S, Khamis HJ, Paliy O: **Distal gut microbiota of adolescent children is different from that of adults**. *Fems Microbiol Ecol* 2011, **77**(2):404-412.
263. Ferrer M, Ruiz A, Lanza F, Haange SB, Oberbach A, Till H, Bargiela R, Campoy C, Segura MT, Richter M *et al*: **Microbiota from the distal guts of lean and obese adolescents exhibit partial functional redundancy besides clear differences in community structure**. *Environ Microbiol* 2013, **15**(1):211-226.
264. Lin A, Bik EM, Costello EK, Dethlefsen L, Haque R, Relman DA, Singh U: **Distinct distal gut microbiome diversity and composition in healthy children from Bangladesh and the United States**. *Plos One* 2013, **8**(1):e53838.
265. Payne AN, Chassard C, Banz Y, Lacroix C: **The composition and metabolic activity of child gut microbiota demonstrate differential adaptation to varied nutrient loads in an in vitro model of colonic fermentation**. *Fems Microbiol Ecol* 2012, **80**(3):608-623.
266. Salonen A, Salojarvi J, Lahti L, de Vos WM: **The adult intestinal core microbiota is determined by analysis depth and health status**. *Clin Microbiol Infec* 2012, **18**:16-20.

267. Mueller S, Saunier K, Hanisch C, Norin E, Alm L, Midtvedt T, Cresci A, Silvi S, Orpianesi C, Verdenelli MC *et al*: **Differences in fecal microbiota in different European study populations in relation to age, gender, and country: a cross-sectional study**. *Appl Environ Microbiol* 2006, **72**(2):1027-1033.
268. Kim SC, Tonkonogy SL, Karrasch T, Jobin C, Sartor RB: **Dual-association of gnotobiotic IL-10-/- mice with 2 nonpathogenic commensal bacteria induces aggressive pancolitis**. *Inflammatory bowel diseases* 2007, **13**(12):1457-1466.
269. Rooks MG, Veiga P, Wardwell-Scott LH, Tickle T, Segata N, Michaud M, Gallini CA, Beal C, van Hylckama-Vlieg JE, Ballal SA *et al*: **Gut microbiome composition and function in experimental colitis during active disease and treatment-induced remission**. *The ISME journal* 2014, **8**(7):1403-1417.
270. Kohane IS, McMurry A, Weber G, MacFadden D, Rappaport L, Kunkel L, Bickel J, Wattanasin N, Spence S, Murphy S *et al*: **The co-morbidity burden of children and young adults with autism spectrum disorders**. *Plos One* 2012, **7**(4):e33224.
271. Mutlu EA, Gillevet PM, Rangwala H, Sikaroodi M, Naqvi A, Engen PA, Kwasny M, Lau CK, Keshavarzian A: **Colonic microbiome is altered in alcoholism**. *Am J Physiol-Gastr L* 2012, **302**(9):G966-G978.
272. Poxton IR, Brown R, Sawyerr A, Ferguson A: **Mucosa-associated bacterial flora of the human colon**. *J Med Microbiol* 1997, **46**(1):85-91.
273. Zoetendal EG, von Wright A, Vilpponen-Salmela T, Ben-Amor K, Akkermans ADL, de Vos WM: **Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces**. *Appl Environ Microb* 2002, **68**(7):3401-3407.
274. Wang M, Ahrne S, Jeppsson B, Molin G: **Comparison of bacterial diversity along the human intestinal tract by direct cloning and sequencing of 16S rRNA genes**. *Fems Microbiol Ecol* 2005, **54**(2):219-231.
275. Lepage P, Seksik P, Sutren M, de la Cochetiere MF, Jian R, Marteau P, Dore J: **Biodiversity of the mucosa-associated microbiota is stable along the distal digestive tract in healthy individuals and patients with IBD**. *Inflammatory bowel diseases* 2005, **11**(5):473-480.
276. Green GL, Brostoff J, Hudspith B, Michael M, Mylonaki M, Rayment N, Staines N, Sanderson J, Rampton DS, Bruce KD: **Molecular characterization of the bacteria adherent to human colorectal mucosa**. *Journal of Applied Microbiology* 2006, **100**(3):460-469.
277. Walker AW, Sanderson JD, Churcher C, Parkes GC, Hudspith BN, Rayment N, Brostoff J, Parkhill J, Dougan G, Petrovska L: **High-throughput clone library analysis of the mucosa-associated microbiota reveals dysbiosis and differences between inflamed and non-inflamed regions of the intestine in inflammatory bowel disease**. *Bmc Microbiology* 2011, **11**.
278. Carroll IM, Chang YH, Park J, Sartor RB, Ringel Y: **Luminal and mucosal-associated intestinal microbiota in patients with diarrhea-predominant irritable bowel syndrome**. *Gut Pathog* 2010, **2**(1):19.

279. Peterson DA, Frank DN, Pace NR, Gordon JI: **Metagenomic approaches for defining the pathogenesis of inflammatory bowel diseases**. *Cell host & microbe* 2008, **3**(6):417-427.
280. Sepehri S, Kotlowski R, Bernstein CN, Krause DO: **Microbial diversity of inflamed and noninflamed gut biopsy tissues in inflammatory bowel disease**. *Inflammatory bowel diseases* 2007, **13**(6):675-683.
281. Gophna U, Sommerfeld K, Gophna S, Doolittle WF, Veldhuyzen van Zanten SJ: **Differences between tissue-associated intestinal microfloras of patients with Crohn's disease and ulcerative colitis**. *J Clin Microbiol* 2006, **44**(11):4136-4141.
282. Zitomersky NL, Atkinson BJ, Franklin SW, Mitchell PD, Snapper SB, Comstock LE, Bousvaros A: **Characterization of Adherent Bacteroidales from Intestinal Biopsies of Children and Young Adults with Inflammatory Bowel Disease**. *Plos One* 2013, **8**(6).
283. Bibiloni R, Mangold M, Madsen KL, Fedorak RN, Tannock GW: **The bacteriology of biopsies differs between newly diagnosed, untreated, Crohn's disease and ulcerative colitis patients**. *J Med Microbiol* 2006, **55**(8):1141-1149.
284. Sokol H, Lepage P, Seksik P, Dore J, Marteau P: **Molecular comparison of dominant microbiota associated with injured versus healthy mucosa in ulcerative colitis**. *Gut* 2007, **56**(1):152-154.
285. Vasquez N, Mangin I, Lepage P, Seksik P, Duong JP, Blum S, Schiffrin E, Suau A, Allez M, Vernier G *et al*: **Patchy distribution of mucosal lesions in ileal Crohn's disease is not linked to differences in the dominant mucosa-associated bacteria: a study using fluorescence in situ hybridization and temporal temperature gradient gel electrophoresis**. *Inflammatory bowel diseases* 2007, **13**(6):684-692.
286. Durban A, Abellan JJ, Jimenez-Hernandez N, Ponce M, Ponce J, Sala T, D'Auria G, Latorre A, Moya A: **Assessing gut microbial diversity from feces and rectal mucosa**. *Microbial ecology* 2011, **61**(1):123-133.
287. Sonnenburg JL, Xu J, Leip DD, Chen CH, Westover BP, Weatherford J, Buhler JD, Gordon JI: **Glycan foraging in vivo by an intestine-adapted bacterial symbiont**. *Science* 2005, **307**(5717):1955-1959.
288. Schluter J, Foster KR: **The Evolution of Mutualism in Gut Microbiota Via Host Epithelial Selection**. *Plos Biology* 2012, **10**(11).
289. Mimouna S, Goncalves D, Barnich N, Darfeuille-Michaud A, Hofman P, Vouret-Craviari V: **Crohn disease-associated Escherichia coli promote gastrointestinal inflammatory disorders by activation of HIF-dependent responses**. *Gut microbes* 2011, **2**(6):335-346.
290. Harrison A, Bakaletz LO, Munson RS, Jr.: **Haemophilus influenzae and oxidative stress**. *Frontiers in cellular and infection microbiology* 2012, **2**:40.
291. Zhang Z, Geng J, Tang X, Fan H, Xu J, Wen X, Ma ZS, Shi P: **Spatial heterogeneity and co-occurrence patterns of human mucosal-associated intestinal microbiota**. *The ISME journal* 2014, **8**(4):881-893.
292. Ahmed S, Macfarlane GT, Fite A, McBain AJ, Gilbert P, Macfarlane S: **Mucosa-associated bacterial diversity in relation to human terminal ileum and colonic biopsy samples**. *Appl Environ Microb* 2007, **73**(22):7435-7442.

293. Schjorring S, Krogfelt KA: **Assessment of bacterial antibiotic resistance transfer in the gut**. *International journal of microbiology* 2011, **2011**:312956.
294. Sekirov I, Tam NM, Jogova M, Robertson ML, Li Y, Lupp C, Finlay BB: **Antibiotic-induced perturbations of the intestinal microbiota alter host susceptibility to enteric infection**. *Infection and immunity* 2008, **76**(10):4726-4736.
295. Engevik MA, Aihara E, Montrose MH, Shull GE, Hassett DJ, Worrell RT: **Loss of NHE3 alters gut microbiota composition and influences Bacteroides thetaiotaomicron growth**. *Am J Physiol Gastrointest Liver Physiol* 2013, **305**(10):G697-711.
296. Said HS, Suda W, Nakagome S, Chinen H, Oshima K, Kim S, Kimura R, Iraha A, Ishida H, Fujita J *et al*: **Dysbiosis of salivary microbiota in inflammatory bowel disease and its association with oral immunological biomarkers**. *DNA research : an international journal for rapid publication of reports on genes and genomes* 2014, **21**(1):15-25.
297. Brito F, Zaltman C, Carvalho ATP, Fischer RG, Persson R, Gustafsson A, Figueredo CMS: **Subgingival microflora in inflammatory bowel disease patients with untreated periodontitis**. *Eur J Gastroen Hepat* 2013, **25**(2):239-245.
298. Juste C, Kreil DP, Beauvallet C, Guillot A, Vaca S, Carapito C, Mondot S, Sykacek P, Sokol H, Blon F *et al*: **Bacterial protein signals are associated with Crohn's disease**. *Gut* 2014.
299. Haberman Y, Tickle TL, Dexheimer PJ, Kim MO, Tang D, Karns R, Baldassano RN, Noe JD, Rosh J, Markowitz J *et al*: **Pediatric Crohn disease patients exhibit specific ileal transcriptome and microbiome signature**. *J Clin Invest* 2014.
300. Segata N, Haake SK, Mannon P, Lemon KP, Waldron L, Gevers D, Huttenhower C, Izard J: **Composition of the adult digestive tract bacterial microbiome based on seven mouth surfaces, tonsils, throat and stool samples**. *Genome Biol* 2012, **13**(6):R42.
301. Liu J, Wu C, Huang IH, Merritt J, Qi F: **Differential response of Streptococcus mutans towards friend and foe in mixed-species cultures**. *Microbiology* 2011, **157**(Pt 9):2433-2444.
302. Negroni A, Costanzo M, Vitali R, Superti F, Bertuccini L, Tinari A, Minelli F, Di Nardo G, Nuti F, Pierdomenico M *et al*: **Characterization of adherent-invasive Escherichia coli isolated from pediatric patients with inflammatory bowel disease**. *Inflammatory bowel diseases* 2012, **18**(5):913-924.
303. Kleessen B, Kroesen AJ, Buhr HJ, Blaut M: **Mucosal and invading bacteria in patients with inflammatory bowel disease compared with controls**. *Scand J Gastroenterol* 2002, **37**(9):1034-1041.
304. Karlsson H, Larsson P, Wold AE, Rudin A: **Pattern of cytokine responses to gram-positive and gram-negative commensal bacteria is profoundly changed when monocytes differentiate into dendritic cells**. *Infection and immunity* 2004, **72**(5):2671-2678.
305. Gupta RS: **Origin of diderm (Gram-negative) bacteria: antibiotic selection pressure rather than endosymbiosis likely led to the evolution of bacterial cells with two membranes**. *Anton Leeuw Int J G* 2011, **100**(2):171-182.

306. Tyrer P, Foxwell AR, Cripps AW, Apicella MA, Kyd JM: **Microbial pattern recognition receptors mediate M-cell uptake of a gram-negative bacterium**. *Infection and immunity* 2006, **74**(1):625-631.
307. Washio J, Shimada Y, Yamada M, Sakamaki R, Takahashi N: **Hydrogen Sulfide Production by Oral Veillonella: Effect of pH and Lactate**. *Appl Environ Microbiol* 2014.
308. Shatalin K, Shatalina E, Mironov A, Nudler E: **H2S: a universal defense against antibiotics in bacteria**. *Science* 2011, **334**(6058):986-990.
309. Takeshita T, Suzuki N, Nakano Y, Yasui M, Yoneda M, Shimazaki Y, Hirofuji T, Yamashita Y: **Discrimination of the oral microbiota associated with high hydrogen sulfide and methyl mercaptan production**. *Scientific reports* 2012, **2**:215.
310. Fujimoto T, Imaeda H, Takahashi K, Kasumi E, Bamba S, Fujiyama Y, Andoh A: **Decreased abundance of Faecalibacterium prausnitzii in the gut microbiota of Crohn's disease**. *Journal of gastroenterology and hepatology* 2013, **28**(4):613-619.
311. Yang F, Huang S, He T, Catrenich C, Teng F, Bo C, Chen J, Liu J, Li J, Song Y *et al*: **Microbial basis of oral malodor development in humans**. *Journal of dental research* 2013, **92**(12):1106-1112.
312. Roediger WE, Moore J, Babidge W: **Colonic sulfide in pathogenesis and treatment of ulcerative colitis**. *Dig Dis Sci* 1997, **42**(8):1571-1579.
313. Loubinoux J, Bronowicki JP, Pereira IA, Mougenel JL, Faou AE: **Sulfate-reducing bacteria in human feces and their association with inflammatory bowel diseases**. *Fems Microbiol Ecol* 2002, **40**(2):107-112.
314. Pandey S, Singh A, Kumar P, Chaudhari A, Nareshkumar G: **Probiotic Escherichia coli CFR 16 Producing Pyrroloquinoline Quinone (PQQ) Ameliorates 1,2-Dimethylhydrazine-Induced Oxidative Damage in Colon and Liver of Rats**. *Applied biochemistry and biotechnology* 2014.
315. Smith PM, Howitt MR, Panikov N, Michaud M, Gallini CA, Bohlooly YM, Glickman JN, Garrett WS: **The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis**. *Science* 2013, **341**(6145):569-573.
316. Atarashi K, Tanoue T, Oshima K, Suda W, Nagano Y, Nishikawa H, Fukuda S, Saito T, Narushima S, Hase K *et al*: **Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota**. *Nature* 2013, **500**(7461):232-236.
317. Walter J, Ley R: **The human gut microbiome: ecology and recent evolutionary changes**. *Annual review of microbiology* 2011, **65**:411-429.
318. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T *et al*: **A human gut microbial gene catalogue established by metagenomic sequencing**. *Nature* 2010, **464**(7285):59-65.
319. Rubinstein MR, Wang X, Liu W, Hao Y, Cai G, Han YW: **Fusobacterium nucleatum promotes colorectal carcinogenesis by modulating E-cadherin/beta-catenin signaling via its FadA adhesin**. *Cell host & microbe* 2013, **14**(2):195-206.
320. Strauss J, Kaplan GG, Beck PL, Rioux K, Panaccione R, Devinney R, Lynch T, Allen-Vercoe E: **Invasive potential of gut mucosa-derived Fusobacterium nucleatum positively correlates with IBD status of the host**. *Inflammatory bowel diseases* 2011, **17**(9):1971-1978.

321. Sandkvist M: **Type II secretion and pathogenesis**. *Infection and immunity* 2001, **69**(6):3523-3535.
322. Carroll IM, Maharshak N: **Enteric bacterial proteases in inflammatory bowel disease- pathophysiology and clinical implications**. *World J Gastroenterol* 2013, **19**(43):7531-7543.
323. Carbonero F, Benefiel AC, Alizadeh-Ghamsari AH, Gaskins HR: **Microbial pathways in colonic sulfur metabolism and links with health and disease**. *Frontiers in physiology* 2012, **3**:448.
324. Fite A, Macfarlane S, Furrie E, Bahrami B, Cummings JH, Steinke DT, Macfarlane GT: **Longitudinal analyses of gut mucosal microbiotas in ulcerative colitis in relation to patient age and disease severity and duration**. *J Clin Microbiol* 2013, **51**(3):849-856.
325. Copeland A, Sikorski J, Lapidus A, Nolan M, Del Rio TG, Lucas S, Chen F, Tice H, Pitluck S, Cheng JF *et al*: **Complete genome sequence of Atopobium parvulum type strain (IPP 1246)**. *Standards in genomic sciences* 2009, **1**(2):166-173.
326. Kazor CE, Mitchell PM, Lee AM, Stokes LN, Loesche WJ, Dewhirst FE, Paster BJ: **Diversity of bacterial populations on the tongue dorsa of patients with halitosis and healthy patients**. *J Clin Microbiol* 2003, **41**(2):558-563.
327. Yen D, Cheung J, Scheerens H, Poulet F, McClanahan T, McKenzie B, Kleinschek MA, Owyang A, Mattson J, Blumenschein W *et al*: **IL-23 is essential for T cell-mediated colitis and promotes inflammation via IL-17 and IL-6**. *J Clin Invest* 2006, **116**(5):1310-1316.
328. Knights D, Lassen KG, Xavier RJ: **Advances in inflammatory bowel disease pathogenesis: linking host genetics and the microbiome**. *Gut* 2013, **62**(10):1505-1510.
329. Coccia M, Harrison OJ, Schiering C, Asquith MJ, Becher B, Powrie F, Maloy KJ: **IL-1beta mediates chronic intestinal inflammation by promoting the accumulation of IL-17A secreting innate lymphoid cells and CD4(+) Th17 cells**. *J Exp Med* 2012, **209**(9):1595-1609.
330. Arijs I, Vanhove W, Rutgeerts P, Schuit F, Verbeke K, De Preter V: **Decreased mucosal sulfide detoxification capacity in patients with Crohn's disease**. *Inflammatory bowel diseases* 2013, **19**(5):E70-72.
331. De Preter V, Arijs I, Windey K, Vanhove W, Vermeire S, Schuit F, Rutgeerts P, Verbeke K: **Decreased mucosal sulfide detoxification is related to an impaired butyrate oxidation in ulcerative colitis**. *Inflammatory bowel diseases* 2012, **18**(12):2371-2380.
332. Hildebrandt TM, Grieshaber MK: **Three enzymatic activities catalyze the oxidation of sulfide to thiosulfate in mammalian and invertebrate mitochondria**. *The FEBS journal* 2008, **275**(13):3352-3361.
333. Linden DR: **Hydrogen sulfide signaling in the gastrointestinal tract**. *Antioxid Redox Signal* 2014, **20**(5):818-830.
334. Cooper CE, Brown GC: **The inhibition of mitochondrial cytochrome oxidase by the gases carbon monoxide, nitric oxide, hydrogen cyanide and hydrogen sulfide: chemical mechanism and physiological significance**. *Journal of bioenergetics and biomembranes* 2008, **40**(5):533-539.

335. Roediger WE, Duncan A, Kapaniris O, Millard S: **Reducing sulfur compounds of the colon impair colonocyte nutrition: implications for ulcerative colitis**. *Gastroenterology* 1993, **104**(3):802-809.
336. Kaiserling E: **Newly-formed lymph nodes in the submucosa in chronic inflammatory bowel disease**. *Lymphology* 2001, **34**(1):22-29.
337. Mackay F, Browning JL, Lawton P, Shah SA, Comiskey M, Bhan AK, Mizoguchi E, Terhorst C, Simpson SJ: **Both the lymphotoxin and tumor necrosis factor pathways are involved in experimental murine models of colitis**. *Gastroenterology* 1998, **115**(6):1464-1475.
338. Jiang W, Wang XQ, Zeng BH, Liu L, Tardivel A, Wei H, Han JH, MacDonald HR, Tschopp J, Tian ZG *et al*: **Recognition of gut microbiota by NOD2 is essential for the homeostasis of intestinal intraepithelial lymphocytes**. *J Exp Med* 2013, **210**(11):2465-2476.
339. Salvador JAR, Figueiredo SAC, Pinto RMA, Silvestre SM: **Bismuth compounds in medicinal chemistry**. *Future Med Chem* 2012, **4**(11):1495-1523.
340. Sheridan WG, Lowndes RH, Young HL: **Tissue Oxygen-Tension as a Predictor of Colonic Anastomotic Healing**. *Diseases of the Colon & Rectum* 1987, **30**(11):867-871.
341. Lupp C, Robertson ML, Wickham ME, Sekirov I, Champion OL, Gaynor EC, Finlay BB: **Host-mediated inflammation disrupts the intestinal microbiota and promotes the overgrowth of Enterobacteriaceae**. *Cell host & microbe* 2007, **2**(3):204.
342. Rajilic-Stojanovic M: **Function of the microbiota**. *Best practice & research Clinical gastroenterology* 2013, **27**(1):5-16.
343. Ramasamy S, Singh S, Taniere P, Langman MJ, Eggo MC: **Sulfide-detoxifying enzymes in the human colon are decreased in cancer and upregulated in differentiation**. *Am J Physiol Gastrointest Liver Physiol* 2006, **291**(2):G288-296.
344. Bar F, Bochmann W, Widok A, von Medem K, Pagel R, Hirose M, Yu X, Kalies K, Konig P, Bohm R *et al*: **Mitochondrial gene polymorphisms that protect mice from colitis**. *Gastroenterology* 2013, **145**(5):1055-1063 e1053.
345. Miller TW, Wang EA, Gould S, Stein EV, Kaur S, Lim L, Amarnath S, Fowler DH, Roberts DD: **Hydrogen sulfide is an endogenous potentiator of T cell activation**. *The Journal of biological chemistry* 2012, **287**(6):4211-4221.
346. Vital M, Penton CR, Wang Q, Young VB, Antonopoulos DA, Sogin ML, Morrison HG, Raffals L, Chang EB, Huffnagle GB *et al*: **A gene-targeted approach to investigate the intestinal butyrate-producing bacterial community**. *Microbiome* 2013, **1**(1):8.

Contribution of Collaborators

DR. Chritian Jobin and Dr. Marcus Mühlbauer: The work regarding the assessment of *A. parvulum*-induced inflammation in *Il-10*$^{-/-}$ mice was done in collaboration with Dr. Christian Jobin, Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. Dr. Marcus Mühlbauer from Dr. Christian Jobin's lab undertook the experiments involving *Il-10*$^{-/-}$ mice colonscopy, tissue processing and hitological examination, and assessment of cytokines expression level.

Dr. Daniel Figeys and Cheng-Kang Chiang: The work regarding the assessment of human proteome using mass spectrometry was conducted in collaboration with Dr. Daniel Figeys, Department of Biochemistry, Microbiology and Immunology and Ottawa Institute of Systems Biology, University of Ottawa, Ottawa, ON, Canada. Dr. Cheng-Kang Chiang from Dr. Figeys' lab undertook the experiments involving the proteomic assessment of biopsies using mass spectrometry and by employing SILAC approach.

Dr. Mehrdad Hajibabaei and Dr.Shadi Shokralla: Preparation of the 16S rRNA-V6 library for 454 pyrosequencing and generation of the sequences using 454 Genome Sequencer FLX System was conducted by Dr. Shadi Shokralla from the lab of Dr. Mehrdad Hajibabaei, Biodiversity Institute of Ontario, Department of Integrative Biology, University of Guelph, Guelph, ON, Canada.

Dr. Alain Stintzi: Dr. Alain Stintzi (Department of Biochemistry, Microbiology and Immunology, University of Ottawa), contributed both as a supervisor to the project as a whole, and directly contributed to the data analysis for the proteomic assessment in colonic biopsies, and assisted with the maintenance, euthanasia, and necropsies for DSS-gnotobiotic mice.

Appendix I: Primers used for constructing the 16S rRNA-V6 Illumina library.
Nucleotides in bold represent the barcode sequences.

Primer Name	Primer Sequence (5` – 3`)
MF1	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**ATAGCG**AAACTCAAAKGAATTGACGG
MF2	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**AGGGT**AAACTCAAAKGAATTGACGG
MF3	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**TTCAT**AAACTCAAAKGAATTGACGG-3’
MF4	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**GATCGT**AAACTCAAAKGAATTGACGG
MF5	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**GCCCGT**AAACTCAAAKGAATTGACGG
MF6	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**CTGTC**AAACTCAAAKGAATTGACGG
MF7	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**CACGT**AAACTCAAAKGAATTGACGG
MF8	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**CGTACG**AAACTCAAAKGAATTGACGG
MF9	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**GGAC**AAACTCAAAKGAATTGACGG
MF10	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**TAGA**AAACTCAAAKGAATTGACGG
MF11	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**TCAT**AAACTCAAAKGAATTGACGG
MF12	ACACTCTTTCCCTACACGACGCTCTTCCGATCT**ACTT**AAACTCAAAKGAATTGACGG
MR1	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCT**ATAGCGA**ACGAGCTGACGACARCCATG
MR2	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCT**AGGGTA**ACGAGCTGACGACARCCATG
MR3	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCT**TTCATA**ACGAGCTGACGACARCCATG
MR4	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCT**GATCGTA**ACGAGCTGACGACARCCATG
MR5	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCT**GCCCGTA**ACGAGCTGACGACARCCATG
MR6	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCT**CTGTCA**ACGAGCTGACGACARCCATG
MR7	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCT**CACGTA**ACGAGCTGACGACARCCATG
MR8	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCT**CGTACGA**ACGAGCTGACGACARCCATG
MR9	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCT**GGACA**ACGAGCTGACGACARCCATG
MR10	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCT**TAGAA**ACGAGCTGACGACARCCATG
MR11	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCT**TCAT**ACGAGCTGACGACARCCATG
MR12	CTCGGCATTCCTGCTGAACCGCTCTTCCGATCT**ACTT**ACGAGCTGACGACARCCATG
PCR-FWD1	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
PCR-RVS1	CAAGCAGAAGACGCCATACGAGATCGGTCTCGGCATTCCGTCTGAACCGCTCTTCCGATCT

Appendix II: Characteristics of IBD patients and control subjects employed in the current study. RC = Right Colon; LC=Left Colon; TI=Terminal Ileum; BCOA= Biopsy from inflamed tissue; BCON= Biopsy from non-inflamed tissue; CD = Crohn's disease; UC = Ulcerative Colitis; Illumina = Hiseq2500 sequencing; 454pyro = pyrosequencing Roche 454 GS-FLX; PUCAI = Pediatric Ulcerative Colitis Activity Index and PCDAI = Pediatric Crohn's Disease Activity Index.

#Sample ID	Description	Gender	Age	Site	Visual Appearance	Paris Classification	PCDAI/PUCAI	Experiment
HMC012LC	CD	M	13	LC	normal	A1bL1B1G0	12.5 (Mild)	Illumina
HMC014LC	CD	F	14	LC	inflamed	A1bL3L4aB1G1	37.5 (Moderate)	Illumina
HMC015LC	CD	F	14	LC	normal	A1bL1B2G0	10 (Mild)	Illumina
HMC016LC	CD	M	13	LC	inflamed	A1bL3B1G0	20 (Mild)	Illumina
HMC017LC	CD	M	13	LC	inflamed	A1bL2L4aB1G1	57.5 (Severe)	Illumina
HMC022LC	CD	M	14	LC	normal	A1bL1B1G1	37.5 (Moderate)	Illumina
HMC025LC	CD	F	15	LC	normal	A1bL4bB1G0	20 (Mild)	Illumina
HMC029LC	CD	M	14	LC	inflamed	A1bL3L4aB1G0	32.5 (Moderate)	Illumina
HMC038LC	CD	F	16	LC	inflamed	A1bL2B1pG0	20 (Mild)	Illumina
HMC039LC	CD	F	13	LC	normal	A1bL1L4aB1G1	60 (Severe)	Illumina
HMC041LC	CD	F	15	LC	inflamed	A1bL3B1G0	57.5 (Severe)	Illumina
HMC044LC	CD	F	15	LC	normal	A1bL1B3G1	52.5 (Severe)	Illumina
HMC047LC	CD	M	16	LC	normal	A1bL3L4aB1G1	65 (Severe)	Illumina
HMC049LC	CD	M	12	LC	normal	A1bL1B1G1	40 (Severe)	Illumina
HMC050LC	CD	F	16	LC	normal	A1bL1B1G0	12.5 (Mild)	Illumina
HMC051LC	CD	F	16	LC	normal	A1bL3L4aB1G0	65 (Severe)	Illumina
HMC054LC	CD	N	16	LC	inflamed	A1bL2B1G0	5 (Inactive)	Illumina
HMC061LC	CD	M	9	LC	inflamed	A1aL2B1G0	32.5 (Moderate)	Illumina

HMC062LC	CD	F	15	LC	inflamed	A1bL3L4aB3G0	57.5 (Severe)	Illumina
HMC063LC	CD	M	13	LC	normal	A1bL1L4aB1G1	65 (Severe)	Illumina
HMC065LC	CD	M	16	LC	normal	A1bL1B2B3G0	50 (Severe)	Illumina
HMC068LC	CD	M	17	LC	inflamed	A2L2B1G0	32.5 (Moderate)	Illumina
HMC072LC	CD	M	12	LC	inflamed	A1bL3L4aB1G1	55 (Severe)	Illumina
HMC075LC	CD	M	14	LC	inflamed	A1bL2B1G1	50 (Severe)	Illumina
HMC078LC	CD	F	15	LC	normal	A1bL1B1G0	45 (Severe)	Illumina
HMC079LC	CD	M	16	LC	inflamed	A1bL3B1G0	45 (Severe)	Illumina
HMC080LC	CD	M	15	LC	inflamed	A1bL1L4aB1G0	45 (Severe)	Illumina
HMC081LC	CD	M	11	LC	normal	A1bL1B1G1	67.5 (Severe)	Illumina
HMC082LC	CD	F	15	LC	inflamed	A1bL3B1G0	32.5 (Moderate)	Illumina
HMC083LC	CD	M	14	LC	inflamed	A1bL2L4aB1G1	20 (Mild)	Illumina
HMC084LC	CD	M	16	LC	inflamed	A1bL2B1G0	42.5 (Severe)	Illumina
HMC085LC	CD	M	16	LC	inflamed	A1bL2B1G1	62.5 (Severe)	Illumina
HMC089LC	CD	M	14	LC	inflamed	A1bL3B1G0	52.5 (Severe)	Illumina
HMC090LC	CD	M	16	LC	inflamed	A1bL3L4B1G0	52.5 (Severe)	Illumina
HMC093LC	CD	M	12	LC	normal	A1bL1B1pG0	27.5 (Mild)	Illumina
HMC094LC	CD	M	11	LC	normal	A1bL1B1pG0	45 (Severe)	Illumina
HMC095LC	CD	M	11	LC	normal	A1bL3B1pG1	65 (Severe)	Illumina
HMC097LC	CD	M	12	LC	inflamed	A1bL3L4aB1G0	40 (Severe)	Illumina
HMC118LC	CD	F	6	LC	inflamed	A1aL3B1G0	30 (moderate)	Illumina
HMC012RC	CD	M	13	RC	normal	A1bL1B1G0	12.5 (Mild)	Illumina
HMC014RC	CD	F	14	RC	normal	A1bL3L4aB1G1	37.5 (Moderate)	454pyro
HMC015RC	CD	F	14	RC	normal	A1bL1B2G0	10 (Mild)	Illumina

HMC016RC	CD	M	13	RC	normal	A1bL3B1G0	20 (Mild)	Illumina
HMC017RC	CD	M	13	RC	inflamed	A1bL2L4aB1G1	57.5 (Severe)	Illumina
HMC022RC	CD	M	14	RC	normal	A1bL1B1G1	37.5 (Moderate)	Illumina
HMC025RC	CD	F	15	RC	normal	A1bL4bB1G0	20 (Mild)	Illumina
HMC029RC	CD	M	14	RC	inflamed	A1bL3L4aB1G0	32.5 (Moderate)	Illumina
HMC030RC	CD	F	17	RC	inflamed	A2L3L4aB1G0	45 (Severe)	Illumina
HMC038RC	CD	F	16	RC	inflamed	A1bL2B1pG0	20 (Mild)	Illumina
HMC039RC	CD	F	13	RC	normal	A1bL1L4aB1G1	60 (Severe)	Illumina
HMC041RC	CD	F	15	RC	normal	A1bL3B1G0	57.5 (Severe)	Illumina
HMC044RC	CD	F	15	RC	normal	A1bL1B3G1	52.5 (Severe)	Illumina
HMC047RC	CD	M	16	RC	inflamed	A1bL3L4aB1G1	65 (Severe)	Illumina
HMC049RC	CD	M	12	RC	normal	A1bL1B1G1	40 (Severe)	Illumina/ 454pyro
HMC050RC	CD	F	16	RC	normal	A1bL1B1G0	12.5 (Mild)	Illumina
HMC051RC	CD	F	16	RC	inflamed	A1bL3L4aB1G0	65 (Severe)	Illumina
HMC061RC	CD	M	9	RC	inflamed	A1aL2B1G0	32.5 (Moderate)	Illumina/ 454pyro
HMC062RC	CD	F	15	RC	inflamed	A1bL3L4aB3G0	57.5 (Severe)	Illumina
HMC063RC	CD	M	13	RC	normal	A1bL1L4aB1G1	65 (Severe)	Illumina/ 454pyro
HMC065RC	CD	M	16	RC	normal	A1bL1B2B3G0	50 (Severe)	Illumina
HMC068RC	CD	M	17	RC	inflamed	A2L2B1G0	32.5 (Moderate)	454Pyro
HMC072RC	CD	M	12	RC	inflamed	A1bL3L4aB1G1	55 (Severe)	Illumina
HMC075RC	CD	M	14	RC	inflamed	A1bL2B1G1	50 (Severe)	Illumina
HMC078RC	CD	F	15	RC	normal	A1bL1B1G0	45 (Severe)	Illumina
HMC079RC	CD	M	16	RC	inflamed	A1bL3B1G0	45 (Severe)	Illumina
HMC081RC	CD	M	11	RC	normal	A1bL1B1G1	67.5 (Severe)	Illumina

HMC082RC	CD	F	15	RC	inflamed	A1bL3B1G0	32.5 (Mild)	Illumina
HMC085RC	CD	M	16	RC	normal	A1bL2B1G1	62.5 (Severe)	Illumina/ 454pyro
HMC086RC	CD	M	8	RC	normal	A1aL1L4aL4bB1G0	22.5 (Mild)	Illumina
HMC090RC	CD	M	16	RC	inflamed	A1bL3L4B1G0	52.5 (Severe)	Illumina/ 454pyro
HMC093RC	CD	M	12	RC	normal	A1bL1B1pG0	27.5 (Mild)	Illumina/ 454pyro
HMC094RC	CD	M	11	RC	normal	A1bL1B1pG0	45 (Severe)	Illumina
HMC095RC	CD	M	11	RC	normal	A1bL3B1pG1	65 (Severe)	Illumina
HMC097RC	CD	M	12	RC	inflamed	A1bL3L4aB1G0	40 (Severe)	454Pyro
HMC201RC	CD	F	11	RC	inflamed	A1bL2B1G1	37.5 (Moderate)	Illumina/Masspec/qRTPCR
HMC202RC	CD	M	17	RC	inflamed	A2L3B1G0	37.5 (Moderate)	Illumina/Masspec/qRTPCR
HMC203RC	CD	M	11	RC	inflamed	A1bL3B1G0	45 (Severe)	Illumina/Masspec/qRTPCR
HMC204RC	CD	M	13	RC	normal	A1bL1L4bB1G1	45 (Severe)	Illumina/ qRTPCR
HMC205RC	CD	M	13	RC	normal	A1bL1L4aB1PG1	45 (Severe)	Illumina/ qRTPCR
HMC206RC	CD	M	14	RC	inflamed	A1bL3B1G1	45 (Severe)	qPCR/qRTPCR
HMC210RC	CD	F	17	RC	normal	A2L1B1G0	37.5 (Moderate)	qRTPCR
HMC213RC	CD	M	10	RC	inflamed	A1bL3B1G1	60 (Severe)	MassSpec
HMC217RC	CD	F	14	RC	normal	A1bL1L4aB1G0	55 (Severe)	qRTPCR
HMC219RC	CD	M	14	RC	inflamed	A1bL3B1G1	37.5 (Moderate)	MassSpec/ qRTPCR
HMC220RC	CD	M	13	RC	inflamed	A1bL3B1G0	60 (Severe)	MassSpec/ qRTPCR
HMC223RC	CD	F	9	RC	inflamed	A1aL2B1G1	45 (Severe)	MassSpec/ qRTPCR
HMC227RC	CD	F	13	RC	inflamed	A1bL3B1G0	32.5 (Moderate)	qRTPCR/MassSpec
HMC228RC	CD	M	13	RC	inflamed	A1bL3B1PG0	62.5 (Severe)	MassSpec/ qRTPCR
HMC229RC	CD	M	10	RC	inflamed	A1bL2L4aB1G0	57.5 (Severe)	Illumina/ qRTPCR
HMC229 BCOA	CD	M	10	RC Biopsy	inflamed	A1bL2L4aB1G0	57.5 (Severe)	Illumina

HMC229 BCON	CD	M	10	RC Biopsy	normal	A1bL2L4aB1G0	57.5 (Severe)	Illumina
HMC229 Stool	CD	M	10	stool	-	A1bL2L4aB1G0	57.5 (Severe)	Illumina
HMC230RC	CD	M	14	RC	inflamed	A1bL3L4aB1PG0	52.5 (Severe)	qRTPCR
HMC231RC	CD	F	15	RC	normal	A1bL1B1G0	37.5 (Moderate)	qPCR/qRTPCR
HMC234RC	CD	F	13	RC	inflamed	A1bL3B1G0	52.5 (Severe)	qRTPCR
HMC239RC	CD	M	3	RC	infamed	A1aL3B1pGO	22.5 (Mild)	MassSpec/ qRTPCR
HMC240RC	CD	M	10	RC	normal	A1bL1B1G0	55 (Severe)	qRTPCR
HMC252RC	CD	M	9	RC	inflamed	A1aL3B1G0	37.5 (Moderate)	MassSpec
HMC253RC	CD	F	11	RC	inflamed	A1bL2L4aB1	40 (Severe)	qPCR
HMC254RC	CD	M	12	RC	normal	A1bL1L4aB2B3	27.5 (Mild)	MassSpec
HMC256RC	CD	F	10	RC	inflamed	A1bL2B1	30 (Moderate)	MassSpec
HMC260RC	CD	M	13	RC	inflamed	A1bL2B1p	47.5 (Severe)	Illumina/qPCR
HMC260 BCOA	CD	M	13	RC Biopsy	inflamed	A1bL2B1p	47.5 (Severe)	Illumina
HMC260 BCON	CD	M	13	RC Biopsy	normal	A1bL2B1p	47.5 (Severe)	Illumina
HMC260 Stool	CD	M	13	Stool	-	A1bL2B1p	47.5 (Severe)	Illumina
HMC261RC	CD	F	13	RC	Inflamed	A1bL3L4B1p	32.5 (Moderate)	Illumina/qPCR
HMC261 BCOA	CD	F	13	RC Biopsy	Inflamed	A1bL3L4B1p	32.5 (Moderate)	Illumina
HMC261 BCON	CD	F	13	RC Biopsy	normal	A1bL3L4B1p	32.5 (Moderate)	Illumina/
HMC261 Stool	CD	F	13	Stool	-	A1bL3L4B1p	32.5 (Moderate)	Illumina
HMC262RC	CD	M	15	RC	inflamed	A1bL3L4aB1p	27.5 (mild)	Illumina
HMC262 BCOA	CD	M	15	RC Biopsy	inflamed	A1bL3L4aB1p	27.5 (mild)	Illumina
HMC262 BCON	CD	M	15	RC Biopsy	normal	A1bL3L4aB1p	27.5 (mild)	Illumina
HMC262 Stool	CD	M	15	Stool	-	A1bL3L4aB1p	27.5 (mild)	Illumina
HMC269RC	CD	M	8	RC	normal	A1aL4aL3B1G0	7.5 (Inactive)	qPCR

HMC271RC	CD	M	15	RC	inflamed	A1bL3B1pGO	62.5 (Severe)	qPCR/ MassSpec
HMC272RC	CD	M	12	RC	inflamed	A1bL3B1pG1	50 (Severe)	MassSpec
HMC276RC	CD	M	12	RC	inflamed	A1bL2B1G1	37.5 (Moderate)	qPCR
HMC280RC	CD	M	15	RC	inflamed	A1bL3B1	35 (Moderate)	MassSpec
HMC285RC	CD	F	8	RC	normal	A1aL3B1G0	22.5 (Mild)	qPCR
HMC286RC	CD	M	15	RC	inflamed	A1bL3L4AB1pG0	10 (Mild)	qPCR
HMC288RC	CD	M	14	RC	inflamed	A1bL3B1p	62.5 (Severe)	qPCR/ MassSpec
HMC289RC	CD	M	13	RC	inflamed	A1bL3B1pG0	67.5 (Severe)	qPCR
HMC294RC	CD	M	13	RC	inflamed	A1bL4aL3B1G0	62.5 (Severe)	qPCR
HMC295RC	CD	F	14	RC	inflamed	A1bL2B1G0	20 (Mild)	qPCR
HMC298RC	CD	M	12	RC	Normal	A1bL1B1G0	62.5 (Severe)	qPCR
HMC300RC	CD	M	10	RC	inflamed	A1bL3L4aB1G1	27.5 (Mild)	MassSpec
HMC301RC	CD	F	14	RC	inflamed	A1bL3L4aB1G0	60 (Severe)	qPCR/ MassSpec
HMC305RC	CD	M	4	RC	normal	A1aL3L4aB2pG0	12.5 (Mild)	qPCR
HMC311RC	CD	F	15	RC	Inflamed	A1bL1B1p	32.5 (Moderate)	qPCR
HMC316RC	CD	M	14	RC	inflamed	A1bL3L4aB1	17.5 (Mild)	qPCR/ MassSpec
HMC317RC	CD	M	15	RC	normal	A1bL1L4aB1pG0	0 (Inactive)	qPCR
HMC319RC	CD	M	12	RC	normal	A1bL1B1G1	40 (Severe)	qPCR
HMC327RC	CD	F	10	RC	inflamed	A1bL2B1	25 (Mild)	MassSpec
HMC345RC	CD	M	17	RC	inflamed	A2L3L4aB1G0	57.5 (Severe)	qPCR
HMC347RC	CD	M	15	RC	normal	A1bL3L4aB1GO	35 (Moderate)	qPCR
HMC357RC	CD	F	15	RC	inflamed	A1bL2B1G0	17.5 (Mild)	qPCR
HMC358RC	CD	F	16	RC	normal	A1bL3B1G0	20 (Mild)	qPCR
HMC359RC	CD	M	11	RC	normal	A1bL1B1pG0	37.5 (Moderate)	qPCR

HMC363RC	CD	M	14	RC	normal	A1bL1B1pG0	20 (Mild)	qPCR
HMC374RC	CD	M	15	RC	normal	A1bL3B2G0	37.5 (Moderate)	qPCR
HMC377RC	CD	F	10	RC	inflamed	A1bL2L4aB1G0	37.5 (Moderate)	qPCR
HMC378RC	CD	F	8	RC	inflamed	A1aL3	30 (Moderate)	qPCR
HMC379RC	CD	F	10	RC	normal	A1bL1B1pG0	27.5 (Mild)	qPCR
HMC390RC	CD	F	8	RC	inflamed	A1a	moderate	qPCR
HMC391RC	CD	M	15	RC	inflamed	AlbL1BlpG1	5 (Inactive)	qPCR
HMC392RC	CD	M	15	RC	normal	A1bL3B1pG0	0 (Inactive)	qPCR
HMC394RC	CD	M	16	RC	inflamed	A1bL3B1G0	32.5(Moderate)	qPCR
HMC396RC	CD	M	17	RC	inflamed	A2L1B1G1	2.5(Inactive)	qPCR
HMC399RC	CD	F	14	RC	inflamed	A1bL2B1G0	17.5 (Mild)	qPCR
HMC401RC	CD	M	12	RC	normal	A1bL1B2G0	15 (Mild)	qPCR
HMC012TI	CD	M	13	TI	inflamed	A1bL1B1G0	12.5 (Mild)	Illumina
HMC014TI	CD	F	14	TI	inflamed	A1bL3L4aB1G1	37.5 (Moderate)	Illumina
HMC015TI	CD	F	14	TI	normal	A1bL1B2G0	10 (Mild)	Illumina
HMC030TI	CD	F	17	TI	inflamed	A2L3L4aB1G0	45 (Severe)	Illumina
HMC038TI	CD	F	16	TI	normal	A1bL2B1pG0	20 (Mild)	Illumina
HMC039TI	CD	F	13	TI	normal	A1bL1L4aB1G1	60 (Severe)	Illumina
HMC044TI	CD	F	15	TI	inflamed	A1bL1B3G1	52.5 (Severe)	Illumina
HMC047TI	CD	M	16	TI	inflamed	A1bL3L4aB1G1	65 (Severe)	Illumina
HMC049TI	CD	M	12	TI	inflamed	A1bL1B1G1	40 (Severe)	Illumina
HMC050TI	CD	F	16	TI	inflamed	A1bL1B1G0	12.5 (Mild)	Illumina
HMC051TI	CD	F	16	TI	inflamed	A1bL3L4aB1G0	65 (Severe)	Illumina
HMC061TI	CD	M	9	TI	normal	A1aL2B1G0	32.5 (Moderate)	Illumina

HMC063TI	CD	M	13	TI	inflamed	A1bL1L4aB1G1	65 (Severe)	Illumina
HMC065TI	CD	M	16	TI	normal	A1bL1B2B3G0	50 (Severe)	Illumina
HMC068TI	CD	M	17	TI	normal	A2L2B1G0	32.5 (Moderate)	Illumina
HMC072TI	CD	M	12	TI	inflamed	A1bL3L4aB1G1	55 (Severe)	Illumina
HMC075TI	CD	M	14	TI	normal	A1bL2B1G1	50 (Severe)	Illumina
HMC078TI	CD	F	15	TI	inflamed	A1bL1B1G0	45 (Severe)	Illumina
HMC079TI	CD	M	16	TI	inflamed	A1bL3B1G0	45 (Severe)	Illumina
HMC085TI	CD	M	16	TI	normal	A1bL2B1G1	62.5 (Severe)	Illumina
HMC086TI	CD	M	8	TI	inflamed	A1aL1L4aL4bB1G0	22.5 (Mild)	Illumina
HMC090TI	CD	M	16	TI	inflamed	A1bL3L4B1G0	52.5 (Severe)	Illumina
HMC097TI	CD	M	12	TI	inflamed	A1bL3L4aB1G0	40 (Severe)	Illumina
HMC201TI	CD	F	11	TI	normal	A1bL2B1G1	37.5 (Moderate)	Illumina
HMC203TI	CD	M	11	TI	inflamed	A1bL3B1G0	45 (Severe)	Illumina
HMC204TI	CD	M	13	TI	inflamed	A1bL1L4bB1G1	45 (Severe)	Illumina
HMC205TI	CD	M	13	TI	inflamed	A1bL1L4aB1PG1	45 (Severe)	Illumina
HMC109LC	Control	F	3	DC	normal	-	-	Illumina/qRTPCR
HMC112LC	Control	M	9	DC	normal	-	-	Illumina/qRTPCR
HMC117LC	Control	F	15	DC	normal	-	-	Illumina/qRTPCR
HMC003LC	Control	F	9	LC	normal	-	-	Illumina
HMC004LC	Control	M	12	LC	normal	-	-	Illumina
HMC006LC	Control	M	15	LC	normal	-	-	Illumina
HMC008LC	Control	M	15	LC	normal	-	-	Illumina
HMC009LC	Control	F	17	LC	normal	-	-	Illumina
HMC018LC	Control	M	13	LC	normal	-	-	Illumina

HMC020LC	Control	F	12	LC	normal	-	-	Illumina
HMC021LC	Control	M	16	LC	normal	-	-	Illumina
HMC026LC	Control	F	16	LC	normal	-	-	Illumina
HMC028LC	Control	M	13	LC	normal	-	-	Illumina
HMC031LC	Control	M	13	LC	normal	-	-	Illumina
HMC037LC	Control	F	17	LC	normal	-	-	Illumina
HMC040LC	Control	F	7	LC	normal	-	-	Illumina
HMC042LC	Control	M	17	LC	normal	-	-	Illumina
HMC043LC	Control	F	16	LC	normal	-	-	Illumina
HMC052LC	Control	F	8	LC	normal	-	-	Illumina
HMC053LC	Control	M	9	LC	normal	-	-	Illumina
HMC055LC	Control	M	6	LC	normal	-	-	Illumina
HMC056LC	Control	F	8	LC	normal	-	-	Illumina
HMC059LC	Control	M	14	LC	normal	-	-	Illumina
HMC067LC	Control	F	17	LC	normal	-	-	Illumina
HMC070LC	Control	F	8	LC	normal	-	-	Illumina
HMC073LC	Control	F	16	LC	normal	-	-	Illumina
HMC074LC	Control	F	16	LC	normal	-	-	Illumina
HMC087LC	Control	F	18	LC	normal	-	-	Illumina
HMC098LC	Control	F	16	LC	normal	-	-	Illumina
HMC100LC	Control	M	15	LC	normal	-	-	Illumina
HMC102LC	Control	F	17	LC	normal	-	-	Illumina
HMC105LC	Control	F	18	LC	normal	-	-	Illumina
HMC106LC	Control	F	16	LC	normal	-	-	Illumina

HMC110LC	Control	F	4	LC	normal	-	-	Illumina
HMC003RC	Control	F	9	RC	normal	-	-	Illumina
HMC004RC	Control	M	12	RC	normal	-	-	454Pyro
HMC005RC	Control	F	10	RC	normal	-	-	Illumina
HMC006RC	Control	M	15	RC	normal	-	-	Illumina
HMC018RC	Control	M	13	RC	normal	-	-	Illumina
HMC020RC	Control	F	12	RC	normal	-	-	454Pyro
HMC026RC	Control	F	16	RC	normal	-	-	454Pyro
HMC027RC	Control	F	16	RC	normal	-	-	Illumina
HMC028RC	Control	M	13	RC	normal	-	-	Illumina/ 454pyro
HMC042RC	Control	M	17	RC	normal	-	-	Illumina
HMC043RC	Control	F	16	RC	normal	-	-	Illumina
HMC052RC	Control	F	8	RC	normal	-	-	454Pyro
HMC055RC	Control	M	6	RC	normal	-	-	Illumina/ 454pyro
HMC056RC	Control	F	8	RC	normal	-	-	Illumina
HMC059RC	Control	M	14	RC	normal	-	-	Illumina
HMC067RC	Control	F	17	RC	normal	-	-	Illumina
HMC069RC	Control	F	17	RC	normal	-	-	Illumina
HMC070RC	Control	F	8	RC	normal	-	-	454Pyro
HMC071RC	Control	M	11	RC	normal	-	-	Illumina
HMC073RC	Control	F	16	RC	normal	-	-	Illumina
HMC074RC	Control	F	16	RC	normal	-	-	454Pyro
HMC087RC	Control	F	18	RC	normal	-	-	Illumina/ 454pyro
HMC091RC	Control	M	12	RC	normal	-	-	Illumina

HMC098RC	Control	F	16	RC	normal	-	-	Illumina
HMC100RC	Control	M	15	RC	normal	-	-	Illumina
HMC102RC	Control	F	17	RC	normal	-	-	Illumina
HMC208RC	Control	F	16	RC	normal	-	-	Illumina/qPCR/ qRTPCR
HMC211RC	control	M	13	RC	normal	-	-	qRTPCR
HMC212RC	control	F	15	RC	normal	-	-	qRTPCR
HMC221RC	control	F	9	RC	normal	-	-	qPCR/qRTPCR
HMC222RC	control	M	10	RC	normal	-	-	qPCR/qRTPCR
HMC224RC	control	M	15	RC	normal	-	-	qPCR/qRTPCR/ MassSpec
HMC225RC	control	F	15	RC	normal	-	-	qPCR/qRTPCR
HMC232RC	control	F	15	RC	normal	-	-	qPCR/qRTPCR/ MassSpec
HMC235RC	control	M	11	RC	normal	-	-	qRTPCR
HMC237RC	control	M	14	RC	normal	-	-	MassSpec/ qRTPCR
HMC238RC	Control	M	14	RC	normal	-	-	qPCR/qRTPCR
HMC241RC	Control	F	12	RC	normal	-	-	qRTPCR
HMC245RC	Control	M	16	RC	normal	-	-	qRTPCR
HMC246RC	Control	M	13	RC	normal	-	-	qRTPCR
HMC258RC	Control	M	15	RC	normal	-	-	qPCR/ MassSpec
HMC265RC	Control	M	15	RC	normal	-	-	Illumina
HMC265 BCON	Control	M	15	RC Biopsy	normal	-	-	Illumina
HMC265 Stool	Control	M	15	Stool	normal	-	-	Illumina
HMC270RC	Control	F	17	RC	normal	-	-	qPCR
HMC281RC	Control	F	15	RC	normal	-	-	qPCR/ MassSpec
HMC296RC	Control	M	17	RC	normal	-	-	qPCR

HMC297RC	Control	F	17	RC	normal	-	-	qPCR
HMC302RC	Control	M	18	RC	normal	-	-	qPCR
HMC307RC	Control	F	16	RC	normal	-	-	qPCR/ MassSpec
HMC308RC	Control	M	10	RC	normal	-	-	qPCR
HMC309RC	Control	M	10	RC	normal	-	-	qPCR/ MassSpec
HMC313RC	Control	M	16	RC	normal	-	-	qPCR
HMC315RC	Control	M	16	RC	normal	-	-	qPCR/ MassSpec
HMC321RC	Control	F	16	RC	normal	-	-	qPCR
HMC350RC	Control	M	13	RC	normal	-	-	qPCR
HMC362RC	Control	F	16	RC	normal	-	-	qPCR
HMC371RC	Control	M	7	RC	normal	-	-	qPCR
HMC372RC	Control	M	14	RC	normal	-	-	qPCR
HMC381RC	Control	M	17	RC	normal	-	-	qPCR
HMC397RC	Control	F	10	RC	normal	-	-	qPCR
HMC403RC	Control	F	16	RC	normal	-	-	qPCR
HMC005TI	Control	F	10	TI	normal	-	-	Illumina
HMC008TI	Control	M	15	TI	normal	-	-	Illumina
HMC009TI	Control	F	17	TI	normal	-	-	Illumina
HMC018TI	Control	M	13	TI	normal	-	-	Illumina
HMC020TI	Control	F	12	TI	normal	-	-	Illumina
HMC026TI	Control	F	16	TI	normal	-	-	Illumina
HMC031TI	Control	M	13	TI	normal	-	-	Illumina
HMC032TI	Control	F	16	TI	normal	-	-	Illumina
HMC042TI	Control	M	17	TI	normal	-	-	Illumina

HMC043TI	Control	F	16	TI	normal	-	-	Illumina
HMC052TI	Control	F	8	TI	normal	-	-	Illumina
HMC053TI	Control	M	9	TI	normal	-	-	Illumina
HMC070TI	Control	F	8	TI	normal	-	-	Illumina
HMC073TI	Control	F	16	TI	normal	-	-	Illumina
HMC208TI	Control	F	16	TI	normal	-	-	Illumina
HMC113LC	UC	M	15	DC	inflamed	E2S0	0 (Inactive)	Illumina/qRTPCR
HMC002LC	UC	F	16	LC	normal	E1S0	20 (Mild)	Illumina
HMC013LC	UC	M	12	LC	inflamed	E4S1	65 (Severe)	Illumina
HMC019LC	UC	M	14	LC	inflamed	E4S1	80 (Severe)	Illumina
HMC023LC	UC	M	14	LC	inflamed	E4S0	50 (Moderate)	Illumina
HMC045LC	UC	M	17	LC	inflamed	E3S0	45 (Moderate)	Illumina
HMC046LC	UC	F	4	LC	inflamed	E4S1	65 (Severe)	Illumina
HMC058LC	UC	M	17	LC	inflamed	E3S0	25 (mild)	Illumina
HMC064LC	UC	M	18	LC	inflamed	E4S0	0 ((Inactive)	Illumina
HMC066LC	UC	F	18	LC	inflamed	E4S0	35 (Moderate)	Illumina
HMC076LC	UC	F	17	LC	inflamed	E4S0	55 (Moderate)	Illumina
HMC077LC	UC	M	17	LC	inflamed	E3S0	50 (Moderate)	Illumina
HMC092LC	UC	M	12	LC	inflamed	E3S1	80 (Severe)	Illumina
HMC002RC	UC	F	16	RC	normal	E1S0	20 (Mild)	Illumina
HMC013RC	UC	M	12	RC	inflamed	E4S1	65 (Severe)	Illumina
HMC019RC	UC	M	14	RC	inflamed	E4S1	80 (Severe)	Illumina/ 454pyro
HMC023RC	UC	M	14	RC	inflamed	E4S0	50 (Moderate)	Illumina/ 454pyro
HMC024RC	UC	M	16	RC	normal	E3S0	40 (Moderate)	454pyro

HMC045RC	UC	M	17	RC	normal	E3S0	45 (Moderate)	Illumina
HMC046RC	UC	F	4	RC	inflamed	E4S1	65 (Severe)	Illumina/ 454pyro
HMC064RC	UC	M	18	RC	inflamed	E4S0	0 (Inactive)	Illumina
HMC066RC	UC	F	18	RC	inflamed	E4S0	35 (Moderate)	Illumina/ 454pyro
HMC076RC	UC	F	17	RC	inflamed	E4S0	55 (Moderate)	Illumina/ 454pyro
HMC077RC	UC	M	17	RC	normal	E3S0	50 (Moderate)	Illumina/ 454pyro
HMC088RC	UC	M	16	RC	normal	E1S0	20 (Mild)	Illumina
HMC092RC	UC	M	12	RC	normal	E3S1	80 (Severe)	Illumina/ 454pyro
HMC103RC	UC	F	17	RC	inflamed	E4S0	50 (Moderate)	Illumina
HMC207RC	UC	F	13	RC	inflamed	E4S1	70 (Severe)	Illumina
HMC214RC	UC	M	17	RC	normal	E3S0	35 (Moderate)	qRTPCR
HMC215RC	UC	F	12	RC	inflamed	E4S1	65 (Severe)	qRTPCR
HMC243RC	UC	M	8	RC	normal	E1S0	45 (Moderate)	qRTPCR
HMC247RC	UC	F	13	RC	inflamed	E4S1	80 (Severe)	qRTPCR
HMC249RC	UC	M	14	RC	inflamed	E4S0	45 (Moderate)	qRTPCR
HMC273RC	UC	M	15	RC	Normal	E2S0	40 (Moderate)	qPCR
HMC274RC	UC	M	17	RC	Inflamed	E4S1	70 (Severe)	qPCR
HMC299RC	UC	M	11	RC	Inflamed	E4S0	40 (Moderate)	qPCR
HMC303RC	UC	F	15	RC	inflamed	E4S0	45 (Moderate)	qPCR
HMC348RC	UC	F	17	RC	normal	E3S1	85 (Severe)	qPCR
HMC351RC	UC	M	17	RC	inflamed	E4S0	60 (Moderate)	qPCR
HMC354RC	UC	F	15	RC	inflamed	E4S0	25 (mild)	qPCR
HMC360RC	UC	F	15	RC	inflamed	E4S1	65 (Severe)	qPCR
HMC361RC	UC	F	14	RC	inflamed	E4S1	75 (Severe)	qPCR

HMC369RC	UC	F	17	RC	inflamed	E4S1	75 (Severe)	qPCR
HMC373RC	UC	F	14	RC	inflamed	E4S0	5 (Inactive)	qPCR
HMC382RC	UC	F	13	RC	inflamed	E4S1	65 (Severe)	qPCR
HMC398RC	UC	F	12	RC	inflamed	E4S1	70 (Severe)	qPCR
HMC400RC	UC	M	13	RC	normal	E2S0	60 (Moderate)	qPCR
HMC402RC	UC	F	7	RC	inflamed	E4S0	60 (Moderate)	qPCR
HMC405RC	UC	M	15	RC	inflamed	E4S0	55 (Moderate)	qPCR
HMC002TI	UC	F	16	TI	normal	E1S0	20 (Mild)	Illumina
HMC013TI	UC	M	12	TI	normal	E4S1	65 (Severe)	Illumina
HMC019TI	UC	M	14	TI	normal	E4S1	80 (Severe)	Illumina
HMC023TI	UC	M	14	TI	normal	E4S0	50 (Moderate)	Illumina
HMC024TI	UC	M	16	TI	normal	E3S0	40 (Moderate)	Illumina
HMC045TI	UC	M	17	TI	normal	E3S0	45 (Moderate)	Illumina
HMC058TI	UC	M	17	TI	normal	E3S0	25 (mild)	Illumina
HMC066TI	UC	F	18	TI	normal	E4S0	35 (Moderate)	Illumina
HMC076TI	UC	F	17	TI	normal	E4S0	55 (Moderate)	Illumina
HMC092TI	UC	M	12	TI	normal	E3S1	80 (Severe)	Illumina
HMC207TI	UC	F	13	TI	normal	E4S1	70 (Severe)	Illumina

Appendix III: Taxa that vary significantly in abundance in at least one of the three intestinal locations of non-IBD adolescents (LC; left colon, RC; right colon and TI; terminal ileum). Kruskal-Wallis followed by Bonferroni correction for multiple comparisons test. Bonferroni corrected significance level: 0.0167. Values in bold indicate $p<$ 0.0167. Observations indicate the nubmer of samples from each group employed in the analysis. Mean and std. deviation indicate the mean and std. deviation of 16S reads obtained. *p*|LC, *p*|RC and *p*|TI indicate the *p* values obtained by comparison to the LC, RC and TI, respectively.

Variable	Observ-ations	Mean	Std. deviation	*p*|LC	*p*|RC	*p*|TI
p__Chloroflexi |LC	34	3.07E-06	1.79E-05	1	0.646	0.020
p__Chloroflexi |RC	21	0	0	0.646	1	**0.012**
p__Chloroflexi |TI	15	8.6E-06	3.28E-05	0.020	**0.012**	1
c__Bacilli |LC	34	0.014508	0.049913	1	**0.014**	0.033
c__Bacilli |RC	21	0.014092	0.022807	**0.014**	1	0.959
c__Bacilli |TI	15	0.029154	0.058842	0.033	0.959	1
o__Lactobacillales |LC	34	0.013931	0.049829	1	**0.014**	**0.013**
o__Lactobacillales |RC	21	0.011906	0.021257	**0.014**	1	0.800
o__Lactobacillales |TI	15	0.027891	0.057992	**0.013**	0.800	1
f__211ds20 |LC	34	4.6E-06	1.74E-05	1	0.139	0.093
f__211ds20 |RC	21	9.46E-07	2.61E-06	0.139	1	**0.006**
f__211ds20 |TI	15	7.39E-06	1.23E-05	0.093	**0.006**	1
f__211ds20,g__|LC	34	4.6E-06	1.74E-05	1	0.139	0.093
f__211ds20,g__|RC	21	9.46E-07	2.61E-06	0.139	1	**0.006**
f__211ds20,g__|TI	15	7.39E-06	1.23E-05	0.093	**0.006**	1
f__Bacteroidaceae |LC	34	0.268384	0.146959	1	**0.013**	0.062
f__Bacteroidaceae |RC	21	0.163888	0.152274	**0.013**	1	0.741
f__Bacteroidaceae |TI	15	0.180311	0.179001	0.062	0.741	1
f__Bradyrhizobiaceae,g__ |LC	34	4E-07	1.1E-06	1	0.429	0.054
f__Bradyrhizobiaceae,g__ |RC	21	1.14E-07	3.81E-07	0.429	1	**0.016**
f__Bradyrhizobiaceae,g__ |TI	15	1.23E-06	2.11E-06	0.054	**0.016**	1
f__Carnobacteriaceae |LC	34	0.000196	0.000463	1	**0.012**	0.041
f__Carnobacteriaceae |RC	21	0.000543	0.000881	**0.012**	1	0.841
f__Carnobacteriaceae |TI	15	0.001037	0.002077	0.041	0.841	1
f__Colwelliaceae |LC	34	0	0	1	1.000	**0.010**
f__Colwelliaceae |RC	21	0	0	1.000	1	0.019
f__Colwelliaceae |TI	15	3.43E-07	9.85E-07	**0.010**	0.019	1
f__Colwelliaceae,g__|LC	34	0	0	1	1.000	**0.010**
f__Colwelliaceae,g__|RC	21	0	0	1.000	1	0.019
f__Colwelliaceae,g__|TI	15	3.43E-07	9.85E-07	**0.010**	0.019	1
f__Halothiobacillaceae |LC	34	2.46E-06	9.08E-06	1	0.218	0.077
f__Halothiobacillaceae |RC	21	1.88E-07	8.62E-07	0.218	1	**0.008**
f__Halothiobacillaceae |TI	15	9.8E-06	2.97E-05	0.077	**0.008**	1
f__Hyphomicrobiaceae |LC	34	1.24E-07	7.22E-07	1	0.646	0.020

f__Hyphomicrobiaceae |RC	21	0	0	0.646	1	**0.012**
f__Hyphomicrobiaceae |TI	15	3.09E-06	1.08E-05	0.020	**0.012**	1
f__Nocardiaceae |LC	34	7.38E-08	4.3E-07	1	0.646	0.020
f__Nocardiaceae |RC	21	0	0	0.646	1	**0.012**
f__Nocardiaceae |TI	15	8.93E-07	3E-06	0.020	**0.012**	1
f__Phyllobacteriaceae |LC	34	0	0	1	1.000	**0.010**
f__Phyllobacteriaceae |RC	21	0	0	1.000	1	0.019
f__Phyllobacteriaceae |TI	15	5.07E-07	1.73E-06	**0.010**	0.019	1
f__Rhizobiaceae |LC	34	1.49E-07	8.66E-07	1	0.836	**0.007**
f__Rhizobiaceae |RC	21	7.81E-08	3.58E-07	0.836	1	0.022
f__Rhizobiaceae |TI	15	2E-06	6.05E-06	**0.007**	0.022	1
f__Sanguibacteraceae |LC	34	0	0	1	1.000	**0.010**
f__Sanguibacteraceae |RC	21	0	0	1.000	1	0.019
f__Sanguibacteraceae |TI	15	2.94E-07	9.43E-07	**0.010**	0.019	1
g__Acetobacterium |LC	34	0	0	1	1.000	**0.010**
g__Acetobacterium |RC	21	0	0	1.000	1	0.019
g__Acetobacterium |TI	15	2.07E-06	7.76E-06	**0.010**	0.019	1
g__Acidaminococcus |LC	34	1.71E-06	1E-05	1	0.637	0.022
g__Acidaminococcus |RC	21	0	0	0.637	1	**0.013**
g__Acidaminococcus |TI	15	3.82E-06	1.11E-05	0.022	**0.013**	1
g__Aquamonas |LC	34	9.94E-08	4.04E-07	1	0.922	**0.006**
g__Aquamonas |RC	21	1.52E-07	6.96E-07	0.922	1	**0.009**
g__Aquamonas |TI	15	2.3E-06	6.67E-06	**0.006**	**0.009**	1
g__Azospira |LC	34	0	0	1	1.000	**0.010**
g__Azospira |RC	21	0	0	1.000	1	0.019
g__Azospira |TI	15	8.4E-07	3.01E-06	**0.010**	0.019	1
g__Bacteroides |LC	34	0.268384	0.146959	1	**0.013**	0.062
g__Bacteroides |RC	21	0.163888	0.152274	**0.013**	1	0.741
g__Bacteroides |TI	15	0.180311	0.179001	0.062	0.741	1
g__Bergeriella |LC	34	0	0	1	1.000	**0.010**
g__Bergeriella |RC	21	0	0	1.000	1	0.019
g__Bergeriella |TI	15	2.28E-07	6.73E-07	**0.010**	0.019	1
g__Bergeyella |LC	34	1.74E-07	1.01E-06	1	0.646	0.020
g__Bergeyella |RC	21	0	0	0.646	1	**0.012**
g__Bergeyella |TI	15	9.14E-07	2.68E-06	0.020	**0.012**	1
g__Brachybacterium |LC	34	0	0	1	1.000	**0.010**
g__Brachybacterium |RC	21	0	0	1.000	1	0.019
g__Brachybacterium |TI	15	2.34E-07	7.17E-07	**0.010**	0.019	1
g__Cedecea |LC	34	6.65E-05	0.000321	1	0.152	0.138
g__Cedecea |RC	21	3.17E-06	1.36E-05	0.152	1	**0.011**
g__Cedecea |TI	15	3.97E-05	9.94E-05	0.138	**0.011**	1
g__Comamonas |LC	34	4.94E-08	2.88E-07	1	0.655	0.018
g__Comamonas |RC	21	0	0	0.655	1	**0.011**
g__Comamonas |TI	15	2.07E-05	7.86E-05	0.018	**0.011**	1
g__Devosia |LC	34	9.91E-08	5.78E-07	1	0.646	0.020
g__Devosia |RC	21	0	0	0.646	1	**0.012**

g__Devosia |TI	15	1.73E-06	6.25E-06	0.020	**0.012**	1
g__Dyella |LC	34	0	0	1	1.000	**0.002**
g__Dyella |RC	21	0	0	1.000	1	**0.004**
g__Dyella |TI	15	9.82E-07	3.24E-06	**0.002**	**0.004**	1
g__Enterovibrio |LC	34	5.33E-06	2.32E-05	1	0.535	0.017
g__Enterovibrio |RC	21	7.01E-07	2.47E-06	0.535	1	**0.007**
g__Enterovibrio |TI	15	9.93E-06	2.53E-05	0.017	**0.007**	1
g__Haemophilus |LC	34	1.31E-05	3.16E-05	1	0.549	0.019
g__Haemophilus |RC	21	2.67E-06	6.33E-06	0.549	1	**0.008**
g__Haemophilus |TI	15	0.000116	0.000276	0.019	**0.008**	1
g__Marinobacterium |LC	34	1.23E-06	5.34E-06	1	0.941	**0.006**
g__Marinobacterium |RC	21	8.3E-07	2.47E-06	0.941	1	**0.015**
g__Marinobacterium |TI	15	2.86E-06	8.11E-06	**0.006**	**0.015**	1
g__Massilia |LC	34	0	0	1	1.000	**0.010**
g__Massilia |RC	21	0	0	1.000	1	0.019
g__Massilia |TI	15	4.61E-07	1.25E-06	**0.010**	0.019	1
g__Methylophaga |LC	34	7.44E-08	4.34E-07	1	0.637	0.022
g__Methylophaga |RC	21	0	0	0.637	1	**0.013**
g__Methylophaga |TI	15	1.69E-07	3.51E-07	0.022	**0.013**	1
g__Methylophilus |LC	34	0	0	1	1.000	**0.010**
g__Methylophilus |RC	21	0	0	1.000	1	0.019
g__Methylophilus |TI	15	1.16E-07	3.07E-07	**0.010**	0.019	1
g__Nitrincola |LC	34	0.00014	0.000725	1	0.370	0.043
g__Nitrincola |RC	21	1.32E-05	3.96E-05	0.370	1	**0.010**
g__Nitrincola |TI	15	9.36E-05	0.000221	0.043	**0.010**	1
g__Oceanobacillus |LC	34	1.01E-06	5.55E-06	1	0.545	**0.009**
g__Oceanobacillus |RC	21	8.63E-07	2.69E-06	0.545	1	0.059
g__Oceanobacillus |TI	15	2.14E-06	5.58E-06	**0.009**	0.059	1
g__Paracoccus |LC	34	2.48E-08	1.44E-07	1	0.655	0.018
g__Paracoccus |RC	21	0	0	0.655	1	**0.011**
g__Paracoccus |TI	15	4.25E-06	1.6E-05	0.018	**0.011**	1
g__Pedobacter |LC	34	0	0	1	1.000	**0.010**
g__Pedobacter |RC	21	0	0	1.000	1	0.019
g__Pedobacter |TI	15	3.38E-07	1.09E-06	**0.010**	0.019	1
g__Photobacterium |LC	34	0.001598	0.009158	1	0.341	0.018
g__Photobacterium |RC	21	0.000124	0.000506	0.341	1	**0.003**
g__Photobacterium |TI	15	0.000166	0.000307	0.018	**0.003**	1
g__Polynucleobacter |LC	34	0	0	1	1.000	**0.010**
g__Polynucleobacter |RC	21	0	0	1.000	1	0.019
g__Polynucleobacter |TI	15	1.09E-07	2.87E-07	**0.010**	0.019	1
g__Rhodanobacter |LC	34	0	0	1	1.000	**0.010**
g__Rhodanobacter |RC	21	0	0	1.000	1	0.019
g__Rhodococcus |LC	34	7.38E-08	4.3E-07	1	0.637	0.022
g__Rhodococcus |RC	21	0	0	0.637	1	**0.013**
g__Rhodococcus |TI	15	1.69E-07	3.51E-07	0.022	**0.013**	1
g__Rubellimicrobium |LC	34	0	0	1	1.000	**0.010**

g__Rubellimicrobium |RC	21	0	0	1.000	1	0.019
g__Rubellimicrobium |TI	15	2.32E-07	6.13E-07	**0.010**	0.019	1
g__Solimonas |LC	34	0	0	1	1.000	**0.010**
g__Solimonas |RC	21	0	0	1.000	1	0.019
g__Solimonas |TI	15	3.44E-07	9.89E-07	**0.010**	0.019	1
g__Streptococcus |LC	34	0.005564	0.020132	1	0.030	**0.016**
g__Streptococcus |RC	21	0.009552	0.020925	0.030	1	0.664
g__Streptococcus |TI	15	0.01229	0.022982	**0.016**	0.664	1
g__Tatumella |LC	34	5.49E-05	0.000261	1	0.205	0.032
g__Tatumella |RC	21	6.29E-06	2.41E-05	0.205	1	**0.003**
g__Tatumella |TI	15	2.41E-05	4.26E-05	0.032	**0.003**	1
g__Tepidimonas |LC	34	1.41E-07	5.03E-07	1	0.276	0.074
g__Tepidimonas |RC	21	0	0	0.276	1	**0.011**
g__Tepidimonas |TI	15	2.85E-07	5.23E-07	0.074	**0.011**	1
g__Vagococcus |LC	34	5.16E-05	0.000227	1	**0.016**	0.046
g__Vagococcus |RC	21	7.64E-05	0.000115	**0.016**	1	0.877
g__Vagococcus |TI	15	0.000164	0.000542	0.046	0.877	1
g__Xenorhabdus |LC	34	0	0	1	1.000	**0.010**
g__Xenorhabdus |RC	21	0	0	1.000	1	0.019
g__Xenorhabdus |TI	15	1.16E-07	3.06E-07	**0.010**	0.019	1
g__Zoogloea |LC	34	3.03E-08	1.77E-07	1	0.646	0.020
g__Zoogloea |RC	21	0	0	0.646	1	**0.012**
g__Zoogloea |TI	15	2.57E-06	9.47E-06	0.020	**0.012**	1
f__Sanguibacteraceae,g__ |LC	34	0	0	1	1.000	**0.010**
f__Sanguibacteraceae,g__ |RC	21	0	0	1.000	1	0.019
f__Sanguibacteraceae,g__ |TI	15	2.94E-07	9.43E-07	**0.010**	0.019	1
f__Streptococcaceae |LC	34	0.005566	0.020132	1	0.030	**0.015**
f__Streptococcaceae |RC	21	0.009556	0.020924	0.030	1	0.645
f__Streptococcaceae |TI	15	0.012294	0.02298	**0.015**	0.645	1
o__Actinomycetales,f__ |LC	34	0	0	1	1.000	**0.010**
o__Actinomycetales,f__ |RC	21	0	0	1.000	1	0.019
o__Actinomycetales,f__ |TI	15	1.64E-07	4.64E-07	**0.010**	0.019	1
o__Burkholderiales,f__|LC	34	5.95E-07	1.55E-06	1	0.032	0.410
o__Burkholderiales,f__|RC	21	0	0	0.032	1	**0.012**
o__Burkholderiales,f__|TI	15	1.92E-06	5.39E-06	0.410	**0.012**	1
o__Methylophilales,f__|LC	34	0	0	1	1.000	**0.010**
o__Methylophilales,f__|RC	21	0	0	1.000	1	0.019
o__Methylophilales,f__|TI	15	1.72E-07	4.76E-07	**0.010**	0.019	1
o__Methylophilales,f__ ,g__|LC	34	0	0	1	1.000	**0.010**
o__Methylophilales,f__ ,g__|RC	21	0	0	1.000	1	0.019
o__Methylophilales,f__ ,g__|TI	15	1.72E-07	4.76E-07	**0.010**	0.019	1
o__Actinomycetales,f__,g__ |LC	34	0	0	1	1.000	**0.010**
o__Actinomycetales,f__,g__ |RC	21	0	0	1.000	1	0.019
o__Actinomycetales,f__,g__ |TI	15	1.64E-07	4.64E-07	**0.010**	0.019	1
o__Rhizobiales,f__ |LC	34	1.24E-07	7.22E-07	1	0.863	**0.007**
o__Rhizobiales,f__ |RC	21	3.76E-08	1.72E-07	0.863	1	0.019

o__Rhizobiales,f__ |TI	15	2.04E-06	6.87E-06	**0.007**	0.019	1
o__Rhizobiales,f__,g__ |LC	34	1.24E-07	7.22E-07	1	0.863	**0.007**
o__Rhizobiales,f__,g__ |RC	21	3.76E-08	1.72E-07	0.863	1	0.019
o__Rhizobiales,f__,g__ |TI	15	2.04E-06	6.87E-06	**0.007**	0.019	1
o__Thiotrichales,f__,g__|LC	34	0	0	1	1.000	**0.010**
o__Thiotrichales,f__,g__|RC	21	0	0	1.000	1	0.019
o__Thiotrichales,f__,g__|TI	15	2.83E-07	8.79E-07	**0.010**	0.019	1
o__Thiotrichales,f__|LC	34	0	0	1	1.000	**0.010**
o__Thiotrichales,f__|RC	21	0	0	1.000	1	0.019
o__Thiotrichales,f__|TI	15	2.83E-07	8.79E-07	**0.010**	0.019	1
o__Xanthomonadales,f__|LC	34	0	0	1	1.000	**0.002**
o__Xanthomonadales,f__|RC	21	0	0	1.000	1	**0.004**
o__Xanthomonadales,f__|TI	15	7.35E-07	1.73E-06	**0.002**	**0.004**	1

Appendix IV: Left colon taxa that vary significantly in abundance in at least one of the three age groups of non-IBD adolescents (A;<10 years, B; 11-14 years and C; 15-18 years). Kruskal-Wallis followed by Bonferroni correction for multiple comparisons test. Bonferroni corrected significance level: 0.0167. Values in bold indicate $p< 0.0167$. Observations indicate the number of samples from each group employed in the analysis. Mean and std. deviation indicate the mean and std. deviation of 16S reads obtained. *p*|A, *p*|B and *p*|C indicate the *p* values obtained by comparison to the age groups A, B and C, respectively.

Variable	Observ-ations	Mean	Std. deviation	*p*|A	*p*|B	*p*|C
c__Chloroplast | A	10	4.6E-06	8.27E-06	1	0.017	**0.013**
c__Chloroplast | B	6	0	0	0.017	1	0.592
c__Chloroplast | C	18	3.1E-07	1.18E-06	**0.013**	0.592	1
c__Sphingobacteria | A	10	1.28E-06	2.5E-06	1	0.044	**0.008**
c__Sphingobacteria | B	6	0	0	0.044	1	1.000
c__Sphingobacteria | C	18	0	0	**0.008**	1.000	1
f__Aerococcaceae|g__ | A	10	2.53E-07	8E-07	1	0.048	0.693
f__Aerococcaceae|g__ | B	6	5.45E-07	6.09E-07	0.048	1	**0.013**
f__Aerococcaceae|g__ | C	18	6.39E-08	2.71E-07	0.693	**0.013**	1
f__Clostridiaceae|g__ | A	10	3.93E-05	8.11E-05	1	0.112	0.375
f__Clostridiaceae|g__ | B	6	8.99E-07	1.34E-06	0.112	1	**0.013**
f__Clostridiaceae|g__ | C	18	0.000199	0.000547	0.375	**0.013**	1
f__ClostridialesFamilyXIII.IncertaeSedis | A	10	0.000112	0.000154	1	**0.007**	0.046
f__ClostridialesFamilyXIII.IncertaeSedis | B	6	1.79E-05	4.05E-05	**0.007**	1	0.201
f__ClostridialesFamilyXIII.IncertaeSedis | C	18	3.61E-05	6.08E-05	0.046	0.201	1
f__Flexibacteraceae | A	10	2.84E-07	4.59E-07	1	0.044	**0.008**
f__Flexibacteraceae | B	6	0	0	0.044	1	1.000
f__Flexibacteraceae | C	18	0	0	**0.008**	1.000	1
f__Lactobacillaceae | A	10	0.022239	0.06952	1	**0.004**	0.084
f__Lactobacillaceae | B	6	1.8E-05	1.32E-05	**0.004**	1	0.088
f__Lactobacillaceae | C	18	0.000119	0.000143	0.084	0.088	1
f__Peptococcaceae | A	10	1.01E-05	8.48E-06	1	**0.014**	0.784
f__Peptococcaceae | B	6	1.7E-06	2.26E-06	**0.014**	1	**0.014**
f__Peptococcaceae | C	18	1.87E-05	3.15E-05	0.784	**0.014**	1
f__Peptococcaceae|g__ | A	10	7.56E-06	8.18E-06	1	0.058	0.694
f__Peptococcaceae|g__ | B	6	1.51E-06	2.36E-06	0.058	1	**0.016**
f__Peptococcaceae|g__ | C	18	1.77E-05	3.12E-05	0.694	**0.016**	1
f__Rhodocyclaceae | A	10	3.69E-06	6.08E-06	1	0.574	**0.010**
f__Rhodocyclaceae | B	6	2.23E-06	4.52E-06	0.574	1	0.121
f__Rhodocyclaceae | C	18	6.47E-07	2.36E-06	**0.010**	0.121	1
f__Ruminococcaceae|g__ | A	10	0.049278	0.046531	1	0.300	0.058

f__Ruminococcaceae|g__ | B	6	0.054689	0.042957	0.300	1	**0.006**
f__Ruminococcaceae|g__ | C	18	0.018584	0.029749	0.058	**0.006**	1
g__Actinobacillus | A	10	1.49E-05	2.95E-05	1	0.190	**0.003**
g__Actinobacillus | B	6	3.96E-06	7.5E-06	0.190	1	0.299
g__Anaerococcus | A	10	3.97E-05	6.64E-05	1	**0.000**	**0.003**
g__Anaerococcus | B	6	0	0	**0.000**	1	0.180
g__Anaerococcus | C	18	2.79E-06	5.78E-06	**0.003**	0.180	1
g__Clostridium | A	10	0.0278	0.021786	1	0.437	**0.011**
g__Clostridium | B	6	0.04436	0.039848	0.437	1	0.201
g__Clostridium | C	18	0.082117	0.093489	**0.011**	0.201	1
g__Eubacterium | A	10	3.13E-05	6.38E-05	1	**0.011**	0.330
g__Eubacterium | B	6	7.91E-07	6.43E-07	**0.011**	1	0.049
g__Eubacterium | C	18	9.68E-06	1.56E-05	0.330	0.049	1
g__Lactobacillus | A	10	0.022239	0.06952	1	**0.004**	0.084
g__Lactobacillus | B	6	1.8E-05	1.32E-05	**0.004**	1	0.088
g__Lactobacillus | C	18	0.000119	0.000143	0.084	0.088	1
g__Proteus | A	10	9.43E-07	1.67E-06	1	0.044	**0.008**
g__Proteus | B	6	0	0	0.044	1	1.000
g__Proteus | C	18	0	0	**0.008**	1.000	1
g__Varibaculum | A	10	1.3E-06	2.23E-06	1	0.032	**0.016**
g__Varibaculum | B	6	0	0	0.032	1	0.744
g__Varibaculum | C	18	1.89E-07	8.04E-07	**0.016**	0.744	1
o__Coriobacteriales|f__ | A	10	7.84E-05	0.000223	1	0.067	0.424
o__Coriobacteriales|f__ | B	6	5.15E-05	6.24E-05	0.067	1	**0.007**
o__Coriobacteriales|f__ | C	18	2.73E-06	7.62E-06	0.424	**0.007**	1
o__Coriobacteriales|f__|g__ | A	10	7.84E-05	0.000223	1	0.067	0.424
o__Coriobacteriales|f__|g__ | B	6	5.15E-05	6.24E-05	0.067	1	**0.007**
o__Coriobacteriales|f__|g__ | C	18	2.73E-06	7.62E-06	0.424	**0.007**	1
o__Rhodocyclales | A	10	3.69E-06	6.08E-06	1	0.574	**0.010**
o__Rhodocyclales | B	6	2.23E-06	4.52E-06	0.574	1	0.121
o__Rhodocyclales | C	18	6.47E-07	2.36E-06	**0.010**	0.121	1
o__Sphingobacteriales | A	10	1.28E-06	2.5E-06	1	0.044	**0.008**
o__Sphingobacteriales | B	6	0	0	0.044	1	1.000
o__Sphingobacteriales | C	18	0	0	**0.008**	1.000	1
o__Streptophyta | A	10	4.6E-06	8.27E-06	1	0.017	**0.013**
o__Streptophyta | B	6	0	0	0.017	1	0.592
o__Streptophyta | C	18	3.1E-07	1.18E-06	**0.013**	0.592	1
o__Streptophyta|f__ | A	10	4.6E-06	8.27E-06	1	0.017	**0.013**
o__Streptophyta|f__ | B	6	0	0	0.017	1	0.592
o__Streptophyta|f__ | C	18	3.1E-07	1.18E-06	**0.013**	0.592	1
o__Streptophyta|f__|g__ | A	10	4.6E-06	8.27E-06	1	0.017	**0.013**
o__Streptophyta|f__|g__ | B	6	0	0	0.017	1	0.592
o__Streptophyta|f__|g__ | C	18	3.1E-07	1.18E-06	**0.013**	0.592	1

Appendix V: Right colon taxa that vary significantly in abundance in at least one of the three age groups of non-IBD adolescents (A;<10 year, B; 11-14 year and C; 15-18 year). Kruskal-Wallis followed by Bonferroni correction for multiple comparisons test. Bonferroni corrected significance level: 0.0167. Values in bold indicate $p < 0.0167$. Observations indicate the number of samples from each group employed in the analysis. Mean and std. deviation indicate the mean and std. deviation of 16S reads obtained. *p*|A, *p*|B and *p*|C indicate the *p* values obtained by comparison to the age groups A, B and C, respectively.

Variable	observations	Mean	Std. deviation	*p*|A	*p*|B	*p*|C
c__Alphaproteobacteria | A	4	5.483E-06	2.762E-06	1	**0.003**	0.055
c__Alphaproteobacteria | B	5	4.688E-07	6.954E-07	**0.003**	1	0.107
c__Alphaproteobacteria | C	12	2.285E-06	2.602E-06	0.055	0.107	1
c__Gammaproteobacteria|o__ | A	4	5.653E-06	5.475E-06	1	0.028	**0.011**
c__Gammaproteobacteria|o__ | B	5	3.708E-07	5.164E-07	0.028	1	0.984
c__Gammaproteobacteria|o__ | C	12	2.303E-06	6.733E-06	**0.011**	0.984	1
c__Gammaproteobacteria|o__|f__ | A	4	5.653E-06	5.475E-06	1	0.028	**0.011**
c__Gammaproteobacteria|o__|f__ | B	5	3.708E-07	5.164E-07	0.028	1	0.984
c__Gammaproteobacteria|o__|f__ | C	12	2.303E-06	6.733E-06	**0.011**	0.984	1
f__Aerococcaceae | A	4	3.074E-04	3.821E-04	1	0.556	0.103
f__Aerococcaceae | B	5	1.521E-03	2.297E-03	0.556	1	**0.012**
f__Aerococcaceae | C	12	1.199E-04	2.154E-04	0.103	**0.012**	1
f__ClostridialesFamilyXIII.IncertaeSedis|g__ | A	4	6.551E-05	9.204E-05	1	0.040	**0.004**
f__ClostridialesFamilyXIII.IncertaeSedis|g__ | B	5	1.005E-05	2.189E-05	0.040	1	0.609
f__ClostridialesFamilyXIII.IncertaeSedis|g__ | C	12	5.850E-07	1.186E-06	**0.004**	0.609	1
f__Shewanellaceae | A	4	9.200E-05	1.094E-04	1	**0.009**	**0.008**
f__Shewanellaceae | B	5	4.220E-07	9.436E-07	**0.009**	1	0.671
f__Shewanellaceae | C	12	8.746E-06	2.521E-05	**0.008**	0.671	1
g__Abiotrophia | A	4	8.521E-05	5.329E-05	1	**0.016**	0.018
g__Abiotrophia | B	5	1.786E-05	2.550E-05	**0.016**	1	0.639
g__Abiotrophia | C	12	2.108E-05	3.314E-05	0.018	0.639	1
g__Adlercreutzia | A	4	1.885E-04	3.195E-04	1	0.263	0.162
g__Adlercreutzia | B	5	1.264E-03	1.632E-03	0.263	1	**0.003**
g__Adlercreutzia | C	12	4.080E-05	7.116E-05	0.162	**0.003**	1
g__Aerococcus | A	4	1.677E-04	2.845E-04	1	**0.007**	0.306
g__Aerococcus | B	5	2.894E-06	1.333E-06	**0.007**	1	0.023
g__Aerococcus | C	12	1.893E-05	2.512E-05	0.306	0.023	1
g__Anaerococcus | A	4	1.172E-05	2.193E-05	1	**0.014**	0.031
g__Anaerococcus | B	5	0.000E+0	0.000E+0	**0.014**	1	0.458

g__Anaerococcus | C	12	2.475E-07	5.798E-07	0.031	0.458	1
g__Enterobacter | A	4	5.393E-06	4.399E-06	1	**0.010**	0.032
g__Enterobacter | B	5	5.300E-07	7.495E-07	**0.010**	1	0.352
g__Enterobacter | C	12	3.929E-06	9.005E-06	0.032	0.352	1
g__Enterococcus | A	4	1.353E-04	1.047E-04	1	**0.014**	0.062
g__Enterococcus | B	5	8.929E-06	1.442E-05	**0.014**	1	0.277
g__Enterococcus | C	12	1.895E-03	6.490E-03	0.062	0.277	1
g__Facklamia | A	4	5.414E-05	8.405E-05	1	0.157	0.261
g__Facklamia | B	5	1.499E-03	2.308E-03	0.157	1	**0.003**
g__Facklamia | C	12	7.128E-05	2.092E-04	0.261	**0.003**	1
g__Finegoldia | A	4	1.025E-06	1.506E-06	1	0.214	**0.016**
g__Finegoldia | B	5	2.860E-07	6.395E-07	0.214	1	0.296
g__Finegoldia | C	12	0.000E+00	0.000E+00	**0.016**	0.296	1
g__Hydrogenophaga | A	4	0.000E+00	0.000E+00	1	0.048	1.000
g__Hydrogenophaga | B	5	3.708E-07	5.164E-07	0.048	1	**0.013**
g__Hydrogenophaga | C	12	0.000E+00	0.000E+00	1.000	**0.013**	1
g__Kangiella | A	4	3.618E-06	2.866E-06	1	**0.014**	0.031
g__Kangiella | B	5	0.000E+00	0.000E+00	**0.014**	1	0.458
g__Kangiella | C	12	2.204E-06	6.760E-06	0.031	0.458	1
g__Marinomonas | A	4	4.133E-06	1.772E-06	1	**0.007**	**0.006**
g__Marinomonas | B	5	1.548E-07	3.461E-07	**0.007**	1	0.661
g__Marinomonas | C	12	1.765E-06	5.019E-06	**0.006**	0.661	1
g__Moryella | A	4	1.504E-05	1.763E-05	1	**0.009**	0.123
g__Moryella | B	5	4.688E-07	6.954E-07	**0.009**	1	0.108
g__Moryella | C	12	5.851E-06	8.473E-06	0.123	0.108	1
g__Oribacterium | A	4	1.725E-04	3.025E-04	1	**0.006**	0.097
g__Oribacterium | B	5	7.400E-07	1.030E-06	**0.006**	1	0.098
g__Oribacterium | C	12	1.882E-05	4.010E-05	0.097	0.098	1
g__Porphyromonas | A	4	8.575E-07	1.125E-06	1	0.557	0.110
g__Porphyromonas | B	5	8.928E-07	5.689E-07	0.557	1	**0.013**
g__Porphyromonas | C	12	9.750E-08	3.377E-07	0.110	**0.013**	1
g__Rhodopseudomonas | A	4	6.600E-07	7.911E-07	1	**0.013**	**0.004**
g__Rhodopseudomonas | B	5	0.000E+00	0.000E+00	**0.013**	1	1.000
g__Rhodopseudomonas | C	12	0.000E+00	0.000E+00	**0.004**	1.000	1
g__Shewanella | A	4	9.200E-05	1.094E-04	1	**0.009**	**0.008**
g__Shewanella | B	5	4.220E-07	9.436E-07	**0.009**	1	0.671
g__Shewanella | C	12	8.746E-06	2.521E-05	**0.008**	0.671	1
g__Subdoligranulum | A	4	1.742E-05	3.018E-05	1	0.027	0.779
g__Subdoligranulum | B	5	1.359E-03	1.526E-03	0.027	1	**0.013**
g__Subdoligranulum | C	12	2.058E-04	5.560E-04	0.779	**0.013**	1

g__Tatumella | A	4	5.480E-06	8.535E-06	1	**0.009**	**0.008**
g__Tatumella | B	5	0.000E+00	0.000E+00	**0.009**	1	0.659
g__Tatumella | C	12	9.185E-06	3.182E-05	**0.008**	0.659	1
g__Thermomonas | A	4	5.673E-06	6.494E-06	1	**0.007**	0.055
g__Thermomonas | B	5	1.548E-07	3.461E-07	**0.007**	1	0.184
g__Thermomonas | C	12	2.327E-06	4.286E-06	0.055	0.184	1
o__Chromatiales | A	4	1.553E-06	1.694E-06	1	**0.002**	**0.000**
o__Chromatiales | B	5	0.000E+00	0.000E+00	**0.002**	1	1.000
o__Chromatiales | C	12	0.000E+00	0.000E+00	**0.000**	1.000	1
o__Coriobacteriales|f__ | A	4	1.501E-05	2.769E-05	1	0.130	0.503
o__Coriobacteriales|f__ | B	5	2.180E-04	3.563E-04	0.130	1	**0.008**
o__Coriobacteriales|f__ | C	12	3.021E-04	1.037E-03	0.503	**0.008**	1
o__Coriobacteriales|f__|g__ | A	4	1.501E-05	2.769E-05	1	0.130	0.503
o__Coriobacteriales|f__|g__ | B	5	2.180E-04	3.563E-04	0.130	1	**0.008**
o__Coriobacteriales|f__|g__ | C	12	3.021E-04	1.037E-03	0.503	**0.008**	1

Appendix VI: Terminal Ileum taxa that vary significantly in abundance in at least one of the three age groups of non-IBD adolescents (A;<10 years, B; 11-14 years and C; 15-18 years). Kruskal-Wallis followed by Bonferroni correction for multiple comparisons test. Bonferroni corrected significance level: 0.0167. Values in bold indicate $p< 0.0167$. Observations indicate the number of samples from each group employed in the analysis. Mean and std. deviation indicate the mean and std. deviation of the 16S reads obtained. *p*|A, p|B and *p*|C indicate the *p* values obtained by comparison to the age groups A, B and C, respectively.

Variable	Samples	Mean	Std. deviation	*p* | A	*p* | B	*p* | C
f__Rikenellaceae|g__ | A	4	0.00000103	0.00000150	1	**0.009**	0.231
f__Rikenellaceae|g__ | B	3	0.00020673	0.00017339	**0.009**	1	0.063
f__Rikenellaceae|g__ | C	8	0.00000632	0.00000854	0.231	0.063	1
g__Clostridium | A	4	0.00003810	0.00002577	1	**0.010**	0.235
g__Clostridium | B	3	0.00099304	0.00083390	**0.010**	1	0.065
g__Clostridium | C	8	0.00020586	0.00019853	0.235	0.065	1
g__Odoribacter | A	4	0.00000659	0.00000435	1	0.034	0.891
g__Odoribacter | B	3	0.00356829	0.00298424	0.034	1	**0.012**
g__Odoribacter | C	8	0.00001048	0.00001690	0.891	**0.012**	1
g__Oscillospira | A	4	0.00001329	0.00001074	1	0.019	0.784
g__Oscillospira | B	3	0.00019660	0.00020697	0.019	1	**0.017**
g__Oscillospira | C	8	0.00002516	0.00002314	0.784	**0.017**	1

Appendix VII: Core OTUs (0.75) that vary significantly in abundance in at least one of the three intestinal regions tested in non-IBD adolescents (LC; left colon, RC; right colon and TI; terminal ileum). Kruskal-Wallis followed by Bonferroni correction for multiple comparisons test. Bonferroni corrected significance level: 0.0167. Values in bold indicate $p< 0.0167$. Mean and std. deviation indicate the mean and std. deviation of 16S reads obtained. *p*|LC, *p*|RC and *p*|TI indicate the *p* values obtained by comparison to the LC, RC and TI, respectively.

Variable	Taxonomy	Mean	Std. deviation	*p*|LC	*p*|RC	*p*|TI
89770 | LC	f__Alcaligenaceae	0	0	1	**< 0.0001**	1.000
89770 | RC	f__Alcaligenaceae	1.79339E-05	3.81E-05	**< 0.0001**	1	**< 0.0001**
89770 | TI	f__Alcaligenaceae	0	0	1.000	**< 0.0001**	1
256589 | LC	f__Clostridiales FamilyXI.IncertaeSedis	0	0	1	1.000	**< 0.0001**
256589 | RC	f__Clostridiales FamilyXI.IncertaeSedis	0	0	1.000	1	**< 0.0001**
256589 | TI	f__Clostridiales FamilyXI.IncertaeSedis	1.08437E-05	3.51E-05	**< 0.0001**	**< 0.0001**	1
429887 | LC	f__Enterobacteriaceae	0	0	1	1.000	**< 0.0001**
429887 | RC	f__Enterobacteriaceae	0	0	1.000	1	**< 0.0001**
429887 | TI	f__Enterobacteriaceae	5.81982E-05	0.000174	**< 0.0001**	**< 0.0001**	1
177892 | LC	f__Lachnospiraceae	0	0	1	1.000	**< 0.0001**
177892 | RC	f__Lachnospiraceae	0	0	1.000	1	**< 0.0001**
177892 | TI	f__Lachnospiraceae	3.31142E-05	5.7E-05	**< 0.0001**	**< 0.0001**	1
179319 | LC	f__Lachnospiraceae	0.00015067	0.000451	1	**< 0.0001**	0.530
179319 | RC	f__Lachnospiraceae	0	0	**< 0.0001**	1	**< 0.0001**
179319 | TI	f__Lachnospiraceae	1.91431E-05	3.39E-05	0.530	**< 0.0001**	1
180638 | LC	f__Lachnospiraceae	0	0	1	1.000	**< 0.0001**
180638 | RC	f__Lachnospiraceae	0	0	1.000	1	**< 0.0001**
180638 | TI	f__Lachnospiraceae	3.42497E-05	8.06E-05	**< 0.0001**	**< 0.0001**	1
181170 | LC	f__Lachnospiraceae	0.000163504	0.000454	1	**< 0.0001**	0.333
181170 | RC	f__Lachnospiraceae	0	0	**< 0.0001**	1	**< 0.0001**
181170 | TI	f__Lachnospiraceae	7.80853E-05	0.000162	0.333	**< 0.0001**	1
184577 | LC	f__Lachnospiraceae	1.35467E-05	4.14E-05	1	**< 0.0001**	**< 0.0001**
184577 | RC	f__Lachnospiraceae	0	0	**< 0.0001**	1	1.000
184577 | TI	f__Lachnospiraceae	0	0	**< 0.0001**	1.000	1
188818 | LC	f__Lachnospiraceae	0	0	1	**< 0.0001**	**< 0.0001**
188818 | RC	f__Lachnospiraceae	0.002601578	0.005994	**< 0.0001**	1	0.756
188818 | TI	f__Lachnospiraceae	0.000734735	0.001902	**< 0.0001**	0.756	1
190874 | LC	f__Lachnospiraceae	0	0	1	1.000	**< 0.0001**
190874 | RC	f__Lachnospiraceae	0	0	1.000	1	**< 0.0001**
190874 | TI	f__Lachnospiraceae	3.23197E-06	4.82E-06	**< 0.0001**	**< 0.0001**	1
195336 | LC	f__Lachnospiraceae	0.000791997	0.002828	1	0.426	**0.002**
195336 | RC	f__Lachnospiraceae	0.001516228	0.004435	0.426	1	0.032
195336 | TI	f__Lachnospiraceae	0.000121328	0.000456	**0.002**	0.032	1

204129 | LC	f__Lachnospiraceae	0.000288621	0.000576	1	**0.014**	0.503
204129 | RC	f__Lachnospiraceae	6.72149E-05	0.000117	**0.014**	1	0.162
204129 | TI	f__Lachnospiraceae	8.39978E-05	0.00011	0.503	0.162	1
206865 | LC	f__Lachnospiraceae	0.000317793	0.001413	1	**< 0.0001**	0.264
206865 | RC	f__Lachnospiraceae	0	0	**< 0.0001**	1	**< 0.0001**
206865 | TI	f__Lachnospiraceae	0.000860336	0.002848	0.264	**< 0.0001**	1
212042 | LC	f__Lachnospiraceae	0	0	1	**< 0.0001**	1.000
212042 | RC	f__Lachnospiraceae	0.000285189	0.000971	**< 0.0001**	1	**< 0.0001**
212042 | TI	f__Lachnospiraceae	0	0	1.000	**< 0.0001**	1
212503 | LC	f__Lachnospiraceae	0.000682123	0.001478	1	**0.005**	0.514
212503 | RC	f__Lachnospiraceae	0.000120424	0.000222	**0.005**	1	0.090
212503 | TI	f__Lachnospiraceae	0.000259481	0.00036	0.514	0.090	1
215433 | LC	f__Lachnospiraceae	0	0	1	**< 0.0001**	1.000
215433 | RC	f__Lachnospiraceae	3.39196E-05	8.13E-05	**< 0.0001**	1	**< 0.0001**
215433 | TI	f__Lachnospiraceae	0	0	1.000	**< 0.0001**	1
289113 | LC	f__Lachnospiraceae	0	0	1	**< 0.0001**	1.000
289113 | RC	f__Lachnospiraceae	1.4686E-05	2.73E-05	**< 0.0001**	1	**< 0.0001**
289113 | TI	f__Lachnospiraceae	0	0	1.000	**< 0.0001**	1
293896 | LC	f__Lachnospiraceae	0.00012876	0.000391	1	**< 0.0001**	0.886
293896 | RC	f__Lachnospiraceae	0	0	**< 0.0001**	1	**< 0.0001**
293896 | TI	f__Lachnospiraceae	2.98967E-05	5.32E-05	0.886	**< 0.0001**	1
296516 | LC	f__Lachnospiraceae	0.00019259	0.000781	1	**< 0.0001**	0.644
296516 | RC	f__Lachnospiraceae	0	0	**< 0.0001**	1	**< 0.0001**
296516 | TI	f__Lachnospiraceae	4.66088E-05	0.000103	0.644	**< 0.0001**	1
302693 | LC	f__Lachnospiraceae	2.07369E-05	4.13E-05	1	**< 0.0001**	0.684
302693 | RC	f__Lachnospiraceae	0	0	**< 0.0001**	1	**< 0.0001**
302693 | TI	f__Lachnospiraceae	2.4904E-05	3.7E-05	0.684	**< 0.0001**	1
319438 | LC	f__Lachnospiraceae	0.000101434	0.000224	1	0.334	**< 0.0001**
319438 | RC	f__Lachnospiraceae	7.71803E-05	0.000239	0.334	1	**0.000**
319438 | TI	f__Lachnospiraceae	0	0	**< 0.0001**	**0.000**	1
329241 | LC	f__Lachnospiraceae	0	0	1	**< 0.0001**	1.000
329241 | RC	f__Lachnospiraceae	0.000683906	0.002531	**< 0.0001**	1	**< 0.0001**
329241 | TI	f__Lachnospiraceae	0	0	1.000	**< 0.0001**	1
350240 | LC	f__Lachnospiraceae	0.00013592	0.000479	1	0.571	**< 0.0001**
350240 | RC	f__Lachnospiraceae	0.000826152	0.002998	0.571	1	**< 0.0001**
350240 | TI	f__Lachnospiraceae	0	0	**< 0.0001**	**< 0.0001**	1
355615 | LC	f__Lachnospiraceae	0.000160762	0.000573	1	**< 0.0001**	**< 0.0001**
355615 | RC	f__Lachnospiraceae	0	0	**< 0.0001**	1	1.000
355615 | TI	f__Lachnospiraceae	0	0	**< 0.0001**	1.000	1
58295 | LC	f__Lachnospiraceae	0	0	1	**< 0.0001**	**< 0.0001**
58295 | RC	f__Lachnospiraceae	2.03896E-05	3.34E-05	**< 0.0001**	1	0.771
58295 | TI	f__Lachnospiraceae	2.70138E-05	5.7E-05	**< 0.0001**	0.771	1
66308 | LC	f__Lachnospiraceae	0	0	1	**< 0.0001**	**< 0.0001**
66308 | RC	f__Lachnospiraceae	2.53682E-05	3.74E-05	**< 0.0001**	1	0.716
66308 | TI	f__Lachnospiraceae	2.38075E-05	3.71E-05	**< 0.0001**	0.716	1
98975 | LC	f__Lachnospiraceae	8.00951E-05	0.000253	1	**< 0.0001**	0.609

98975 | RC	f__Lachnospiraceae	0	0	**< 0.0001**	1	**< 0.0001**
98975 | TI	f__Lachnospiraceae	0.000154773	0.000426	0.609	**< 0.0001**	1
168776 | LC	f__Ruminococcaceae	0.00021908	0.000802	1	**< 0.0001**	**< 0.0001**
168776 | RC	f__Ruminococcaceae	0	0	**< 0.0001**	1	1.000
168776 | TI	f__Ruminococcaceae	0	0	**< 0.0001**	1.000	1
190646 | LC	f__Ruminococcaceae	0.000548648	0.002109	1	**< 0.0001**	0.921
190646 | RC	f__Ruminococcaceae	0	0	**< 0.0001**	1	**< 0.0001**
190646 | TI	f__Ruminococcaceae	3.41766E-05	8.47E-05	0.921	**< 0.0001**	1
206906 | LC	f__Ruminococcaceae	0.000212327	0.000824	1	**< 0.0001**	**< 0.0001**
206906 | RC	f__Ruminococcaceae	0	0	**< 0.0001**	1	1.000
206906 | TI	f__Ruminococcaceae	0	0	**< 0.0001**	1.000	1
298739 | LC	f__Ruminococcaceae	0.000105096	0.000282	1	**< 0.0001**	**< 0.0001**
298739 | RC	f__Ruminococcaceae	0	0	**< 0.0001**	1	1.000
298739 | TI	f__Ruminococcaceae	0	0	**< 0.0001**	1.000	1
59566 | LC	f__Ruminococcaceae	0.000699348	0.001921	1	0.017	**0.011**
59566 | RC	f__Ruminococcaceae	0.001503988	0.002481	0.017	1	0.707
59566 | TI	f__Ruminococcaceae	0.002388128	0.005072	**0.011**	0.707	1
79234 | LC	f__Ruminococcaceae	0	0	1	**< 0.0001**	**< 0.0001**
79234 | RC	f__Ruminococcaceae	2.2368E-05	2.85E-05	**< 0.0001**	1	0.955
79234 | TI	f__Ruminococcaceae	4.25775E-05	0.000118	**< 0.0001**	0.955	1
535825 | LC	f__Veillonellaceae	0	0	1	**< 0.0001**	**< 0.0001**
535825 | RC	f__Veillonellaceae	0.000150762	0.000633	**< 0.0001**	1	0.515
535825 | TI	f__Veillonellaceae	0.000123426	0.0003	**< 0.0001**	0.515	1
106166 | LC	g__Actinomyces	1.2203E-05	2.94E-05	1	**0.013**	0.132
106166 | RC	g__Actinomyces	3.8536E-05	5.81E-05	**0.013**	1	0.513
106166 | TI	g__Actinomyces	4.29332E-05	8.53E-05	0.132	0.513	1
173965 | LC	g__Adlercreutzia	4.86988E-05	0.000122	1	**< 0.0001**	**< 0.0001**
173965 | RC	g__Adlercreutzia	0	0	**< 0.0001**	1	1.000
173965 | TI	g__Adlercreutzia	0	0	**< 0.0001**	1.000	1
222721 | LC	g__Anaerostipes	0	0	1	**< 0.0001**	1.000
222721 | RC	g__Anaerostipes	3.49103E-05	0.000102	**< 0.0001**	1	**< 0.0001**
222721 | TI	g__Anaerostipes	0	0	1.000	**< 0.0001**	1
484296 | LC	g__Atopobium	0	0	1	**< 0.0001**	**< 0.0001**
484296 | RC	g__Atopobium	4.01299E-05	6.34E-05	**< 0.0001**	1	0.976
484296 | TI	g__Atopobium	1.35442E-05	2.08E-05	**< 0.0001**	0.976	1
268332 | LC	g__Bacteroides	3.7719E-05	7.54E-05	1	**0.009**	0.071
268332 | RC	g__Bacteroides	1.10275E-05	1.2E-05	**0.009**	1	0.631
268332 | TI	g__Bacteroides	1.96701E-05	2.75E-05	0.071	0.631	1
357582 | LC	g__Bacteroides	0	0	1	**< 0.0001**	1.000
357582 | RC	g__Bacteroides	0.002128923	0.006153	**< 0.0001**	1	**< 0.0001**
357582 | TI	g__Bacteroides	0	0	1.000	**< 0.0001**	1
110075 | LC	g__BD2-13	0.000421117	0.001704	1	**< 0.0001**	0.322
110075 | RC	g__BD2-13	0	0	**< 0.0001**	1	**< 0.0001**
110075 | TI	g__BD2-13	0.000779736	0.002187	0.322	**< 0.0001**	1
177841 | LC	g__Blautia	0	0	1	1.000	**< 0.0001**
177841 | RC	g__Blautia	0	0	1.000	1	**< 0.0001**

177841 | TI	g__Blautia	1.97251E-05	3.42E-05	**< 0.0001**	**< 0.0001**	1
177848 | LC	g__Blautia	0	0	1	1.000	**< 0.0001**
177848 | RC	g__Blautia	0	0	1.000	1	**< 0.0001**
177848 | TI	g__Blautia	2.65975E-06	2.55E-06	**< 0.0001**	**< 0.0001**	1
180401 | LC	g__Blautia	0.000124295	0.000363	1	0.579	**< 0.0001**
180401 | RC	g__Blautia	7.25109E-05	9.83E-05	0.579	1	**< 0.0001**
180401 | TI	g__Blautia	0	0	**< 0.0001**	**< 0.0001**	1
188588 | LC	g__Blautia	0	0	1	1.000	**< 0.0001**
188588 | RC	g__Blautia	0	0	1.000	1	**< 0.0001**
188588 | TI	g__Blautia	9.33373E-06	1.47E-05	**< 0.0001**	**< 0.0001**	1
196007 | LC	g__Blautia	8.82974E-06	1.88E-05	1	**< 0.0001**	0.433
196007 | RC	g__Blautia	0	0	**< 0.0001**	1	**< 0.0001**
196007 | TI	g__Blautia	5.64424E-06	4.28E-06	0.433	**< 0.0001**	1
196986 | LC	g__Blautia	0	0	1	**< 0.0001**	1.000
196986 | RC	g__Blautia	0.000140766	0.000327	**< 0.0001**	1	**< 0.0001**
196986 | TI	g__Blautia	0	0	1.000	**< 0.0001**	1
197208 | LC	g__Blautia	0	0	1	1.000	**< 0.0001**
197208 | RC	g__Blautia	0	0	1.000	1	**< 0.0001**
197208 | TI	g__Blautia	3.69172E-06	4E-06	**< 0.0001**	**< 0.0001**	1
368166 | LC	g__Blautia	0	0	1	**< 0.0001**	1.000
368166 | RC	g__Blautia	3.96602E-05	9.62E-05	**< 0.0001**	1	**< 0.0001**
368166 | TI	g__Blautia	0	0	1.000	**< 0.0001**	1
529400 | LC	g__Blautia	0	0	1	**< 0.0001**	1.000
529400 | RC	g__Blautia	2.38953E-05	3.88E-05	**< 0.0001**	1	**< 0.0001**
529400 | TI	g__Blautia	0	0	1.000	**< 0.0001**	1
536540 | LC	g__Blautia	0	0	1	**< 0.0001**	**< 0.0001**
536540 | RC	g__Blautia	0.000114322	0.000479	**< 0.0001**	1	0.986
536540 | TI	g__Blautia	2.85931E-05	5E-05	**< 0.0001**	0.986	1
544532 | LC	g__Blautia	0	0	1	1.000	**< 0.0001**
544532 | RC	g__Blautia	0	0	1.000	1	**< 0.0001**
544532 | TI	g__Blautia	0.000177705	0.000677	**< 0.0001**	**< 0.0001**	1
551295 | LC	g__Butyrivibrio	0	0	1	1.000	**< 0.0001**
551295 | RC	g__Butyrivibrio	0	0	1.000	1	**< 0.0001**
551295 | TI	g__Butyrivibrio	0.000676119	0.001815	**< 0.0001**	**< 0.0001**	1
148840 | LC	g__Carnobacterium	0	0	1	**< 0.0001**	**< 0.0001**
148840 | RC	g__Carnobacterium	6.22147E-05	0.000141	**< 0.0001**	1	0.307
148840 | TI	g__Carnobacterium	8.47249E-05	0.000189	**< 0.0001**	0.307	1
258165 | LC	g__Clostridium	1.99148E-05	4.04E-05	1	**< 0.0001**	0.687
258165 | RC	g__Clostridium	0	0	**< 0.0001**	1	**< 0.0001**
258165 | TI	g__Clostridium	2.91167E-05	7.85E-05	0.687	**< 0.0001**	1
291561 | LC	g__Clostridium	0.000173534	0.000519	1	**< 0.0001**	0.357
291561 | RC	g__Clostridium	0	0	**< 0.0001**	1	**< 0.0001**
291561 | TI	g__Clostridium	0.000127025	0.000242	0.357	**< 0.0001**	1
294710 | LC	g__Clostridium	8.13706E-05	0.000209	1	**< 0.0001**	0.760
294710 | RC	g__Clostridium	0	0	**< 0.0001**	1	**< 0.0001**
294710 | TI	g__Clostridium	0.000896102	0.003285	0.760	**< 0.0001**	1

369227 | LC	g__Clostridium	0	0	1	1.000	**< 0.0001**
369227 | RC	g__Clostridium	0	0	1.000	1	**< 0.0001**
369227 | TI	g__Clostridium	4.67458E-06	5.85E-06	**< 0.0001**	**< 0.0001**	1
470236 | LC	g__Clostridium	0	0	1	**< 0.0001**	**< 0.0001**
470236 | RC	g__Clostridium	0.004757054	0.008495	**< 0.0001**	1	0.957
470236 | TI	g__Clostridium	0.004639441	0.007838	**< 0.0001**	0.957	1
563572 | LC	g__Clostridium	0	0	1	**< 0.0001**	1.000
563572 | RC	g__Clostridium	4.3859E-05	8.71E-05	**< 0.0001**	1	**< 0.0001**
563572 | TI	g__Clostridium	0	0	1.000	**< 0.0001**	1
581658 | LC	g__Clostridium	0.003085647	0.00952	1	**< 0.0001**	**< 0.0001**
581658 | RC	g__Clostridium	0	0	**< 0.0001**	1	1.000
581658 | TI	g__Clostridium	0	0	**< 0.0001**	1.000	1
583974 | LC	g__Clostridium	0.000719693	0.003329	1	**< 0.0001**	0.913
583974 | RC	g__Clostridium	0	0	**< 0.0001**	1	**< 0.0001**
583974 | TI	g__Clostridium	8.66004E-05	0.000215	0.913	**< 0.0001**	1
95500 | LC	g__Clostridium	0	0	1	**< 0.0001**	**< 0.0001**
95500 | RC	g__Clostridium	4.31871E-05	0.000127	**< 0.0001**	1	0.676
95500 | TI	g__Clostridium	0.000210639	0.000506	**< 0.0001**	0.676	1
101237 | LC	g__Coprococcus	0.000	0.000	1	**< 0.0001**	**< 0.0001**
101237 | RC	g__Coprococcus	0.000151115	0.000643	**< 0.0001**	1	0.668
101237 | TI	g__Coprococcus	2.00896E-05	3.8E-05	**< 0.0001**	0.668	1
178146 | LC	g__Coprococcus	3.43045E-05	8.66E-05	1	**< 0.0001**	0.787
178146 | RC	g__Coprococcus	0	0	**< 0.0001**	1	**< 0.0001**
178146 | TI	g__Coprococcus	8.41117E-06	1.8E-05	0.787	**< 0.0001**	1
182054 | LC	g__Coprococcus	2.39201E-05	5.08E-05	1	0.405	**< 0.0001**
182054 | RC	g__Coprococcus	6.69523E-06	9.95E-06	0.405	1	**0.000**
182054 | TI	g__Coprococcus	0	0	**< 0.0001**	**0.000**	1
369224 | LC	g__Coprococcus	0.000156382	0.000505	1	**< 0.0001**	0.809
369224 | RC	g__Coprococcus	0	0	**< 0.0001**	1	**< 0.0001**
369224 | TI	g__Coprococcus	6.69891E-05	0.000139	0.809	**< 0.0001**	1
520275 | LC	g__Coprococcus	0	0	1	**< 0.0001**	1.000
520275 | RC	g__Coprococcus	7.36026E-05	0.000189	**< 0.0001**	1	**< 0.0001**
520275 | TI	g__Coprococcus	0	0	1.000	**< 0.0001**	1
180983 | LC	g__Eubacterium	4.76876E-05	7.56E-05	1	0.699	**< 0.0001**
180983 | RC	g__Eubacterium	6.12624E-05	7.84E-05	0.699	1	**< 0.0001**
180983 | TI	g__Eubacterium	0	0	**< 0.0001**	**< 0.0001**	1
199293 | LC	g__Faecalibacterium	0	0	1	1.000	**< 0.0001**
199293 | RC	g__Faecalibacterium	0	0	1.000	1	**< 0.0001**
199293 | TI	g__Faecalibacterium	5.71716E-05	0.000114	**< 0.0001**	**< 0.0001**	1
58296 | LC	g__Faecalibacterium	0	0	1	**< 0.0001**	1.000
58296 | RC	g__Faecalibacterium	7.7445E-06	1.36E-05	**< 0.0001**	1	**< 0.0001**
58296 | TI	g__Faecalibacterium	0	0	1.000	**< 0.0001**	1
341133 | LC	g__Ferrimonas	0	0	1	**< 0.0001**	**< 0.0001**
341133 | RC	g__Ferrimonas	0.000702285	0.002955	**< 0.0001**	1	0.187
341133 | TI	g__Ferrimonas	0.000741159	0.001573	**< 0.0001**	0.187	1
203620 | LC	g__Lachnospira	0	0	1	1.000	**< 0.0001**

203620 | RC	g__Lachnospira	0	0	1.000	1	**< 0.0001**
203620 | TI	g__Lachnospira	2.73763E-05	6.07E-05	**< 0.0001**	**< 0.0001**	1
351753 | LC	g__Lachnospira	0.001514071	0.004104	1	**< 0.0001**	0.817
351753 | RC	g__Lachnospira	0	0	**< 0.0001**	1	**< 0.0001**
351753 | TI	g__Lachnospira	0.003802386	0.013125	0.817	**< 0.0001**	1
192215 | LC	g__Leclercia	0	0	1	1.000	**< 0.0001**
192215 | RC	g__Leclercia	0	0	1.000	1	**< 0.0001**
192215 | TI	g__Leclercia	0.00123002	0.004657	**< 0.0001**	**< 0.0001**	1
13978 | LC	g__Mogibacterium	0.000140765	0.000432	1	**< 0.0001**	**< 0.0001**
13978 | RC	g__Mogibacterium	0	0	**< 0.0001**	1	1.000
13978 | TI	g__Mogibacterium	0	0	**< 0.0001**	1.000	1
102362 | LC	g__Nitrincola	0	0	1	1.000	**< 0.0001**
102362 | RC	g__Nitrincola	0	0	1.000	1	**< 0.0001**
102362 | TI	g__Nitrincola	9.35563E-05	0.000221	**< 0.0001**	**< 0.0001**	1
208485 | LC	g__Oscillospira	0.000109066	0.000302	1	**< 0.0001**	0.531
208485 | RC	g__Oscillospira	0	0	**< 0.0001**	1	**0.000**
208485 | TI	g__Oscillospira	1.91874E-05	5.55E-05	0.531	**0.000**	1
306091 | LC	g__Oscillospira	8.72495E-05	0.000189	1	**< 0.0001**	**< 0.0001**
306091 | RC	g__Oscillospira	0	0	**< 0.0001**	1	1.000
306091 | TI	g__Oscillospira	0	0	**< 0.0001**	1.000	1
246316 | LC	g__Paenibacillus	0	0	1	1.000	**< 0.0001**
246316 | RC	g__Paenibacillus	0	0	1.000	1	**< 0.0001**
246316 | TI	g__Paenibacillus	6.24101E-06	9.14E-06	**< 0.0001**	**< 0.0001**	1
515620 | LC	g__Peptostreptococcus	0	0	1	**< 0.0001**	**< 0.0001**
515620 | RC	g__Peptostreptococcus	4.37412E-05	0.000119	**< 0.0001**	1	0.242
515620 | TI	g__Peptostreptococcus	9.28826E-05	0.000257	**< 0.0001**	0.242	1
217410 | LC	g__Pseudomonas	0	0	1	**< 0.0001**	1.000
217410 | RC	g__Pseudomonas	4.85823E-05	9.67E-05	**< 0.0001**	1	**< 0.0001**
217410 | TI	g__Pseudomonas	0	0	1.000	**< 0.0001**	1
190597 | LC	g__Roseburia	0	0	1	**< 0.0001**	1.000
190597 | RC	g__Roseburia	4.98132E-05	7.51E-05	**< 0.0001**	1	**< 0.0001**
190597 | TI	g__Roseburia	0	0	1.000	**< 0.0001**	1
177911 | LC	g__Ruminococcus	4.02694E-05	7.45E-05	1	0.716	**< 0.0001**
177911 | RC	g__Ruminococcus	2.57858E-05	5.39E-05	0.716	1	**< 0.0001**
177911 | TI	g__Ruminococcus	0	0	**< 0.0001**	**< 0.0001**	1
229097 | LC	g__Ruminococcus	2.06955E-05	5.19E-05	1	**< 0.0001**	0.654
229097 | RC	g__Ruminococcus	0	0	**< 0.0001**	1	**< 0.0001**
229097 | TI	g__Ruminococcus	1.9111E-05	3.99E-05	0.654	**< 0.0001**	1
297182 | LC	g__Ruminococcus	3.11483E-05	6.55E-05	1	**< 0.0001**	**< 0.0001**
297182 | RC	g__Ruminococcus	0	0	**< 0.0001**	1	1.000
297182 | TI	g__Ruminococcus	0	0	**< 0.0001**	1.000	1
361809 | LC	g__Ruminococcus	6.23049E-05	8.22E-05	1	0.515	**< 0.0001**
361809 | RC	g__Ruminococcus	0.000102849	0.000135	0.515	1	**< 0.0001**
361809 | TI	g__Ruminococcus	0	0	**< 0.0001**	**< 0.0001**	1
356733 | LC	g__Staphylococcus	0.000187336	0.000383	1	0.560	**< 0.0001**
356733 | RC	g__Staphylococcus	0.000693574	0.001828	0.560	1	**< 0.0001**

356733 | TI	g__Staphylococcus	0	0	**< 0.0001**	**< 0.0001**	1
262794 | LC	g__Subdoligranulum	0	0	1	**< 0.0001**	1.000
262794 | RC	g__Subdoligranulum	0.000444532	0.000957	**< 0.0001**	1	**< 0.0001**
262794 | TI	g__Subdoligranulum	0	0	1.000	**< 0.0001**	1
570199 | LC	g__Vagococcus	0	0	1	**< 0.0001**	**< 0.0001**
570199 | RC	g__Vagococcus	7.55151E-05	0.000114	**< 0.0001**	1	0.782
570199 | TI	g__Vagococcus	0.000162174	0.00054	**< 0.0001**	0.782	1
190727 | LC	g__Xanthomonas	0	0	1	**< 0.0001**	1.000
190727 | RC	g__Xanthomonas	2.52996E-05	6.59E-05	**< 0.0001**	1	**< 0.0001**
190727 | TI	g__Xanthomonas	0	0	1.000	**< 0.0001**	1
185187 | LC	o__Bacteroidales	0	0	1	**< 0.0001**	1.000
185187 | RC	o__Bacteroidales	0.000390335	0.001685	**< 0.0001**	1	**< 0.0001**
185187 | TI	o__Bacteroidales	0	0	1.000	**< 0.0001**	1
322991 | LC	o__Bacteroidales	0	0	1	**< 0.0001**	1.000
322991 | RC	o__Bacteroidales	1.04088E-05	2.18E-05	**< 0.0001**	1	**< 0.0001**
322991 | TI	o__Bacteroidales	0	0	1.000	**< 0.0001**	1
46789 | LC	o__Bacteroidales	0	0	1	**< 0.0001**	**< 0.0001**
46789 | RC	o__Bacteroidales	0.012128986	0.038686	**< 0.0001**	1	0.749
46789 | TI	o__Bacteroidales	0.006099852	0.015478	**< 0.0001**	0.749	1
212304 | LC	o__Clostridiales	0	0	1	**< 0.0001**	1.000
212304 | RC	o__Clostridiales	7.744E-05	0.000196	**< 0.0001**	1	**< 0.0001**
212304 | TI	o__Clostridiales	0	0	1.000	**< 0.0001**	1
15716 | LC	s__Actinomycesoris	0.001282354	0.003149	1	**< 0.0001**	0.828
15716 | RC	s__Actinomycesoris	0	0	**< 0.0001**	1	**< 0.0001**
15716 | TI	s__Actinomycesoris	0.001084483	0.003244	0.828	**< 0.0001**	1
262319 | LC	s__Bacteroidesstercoris	0	0	1	1.000	**< 0.0001**
262319 | RC	s__Bacteroidesstercoris	0	0	1.000	1	**< 0.0001**
262319 | TI	s__Bacteroidesstercoris	0.000559959	0.001886	**< 0.0001**	**< 0.0001**	1
304393 | LC	s__Blautia producta	0	0	1	**< 0.0001**	**< 0.0001**
304393 | RC	s__Blautia producta	1.51785E-05	2.94E-05	**< 0.0001**	1	0.695
304393 | TI	s__Blautia producta	1.26133E-05	1.64E-05	**< 0.0001**	0.695	1
201790 | LC	s__Enterobacter hormaechei	0	0	1	1.000	**< 0.0001**
201790 | RC	s__Enterobacter hormaechei	0	0	1.000	1	**< 0.0001**
201790 | TI	s__Enterobacter hormaechei	7.58611E-05	0.000259	**< 0.0001**	**< 0.0001**	1
178529 | LC	s__Eubacterium rectale	0	0	1	**< 0.0001**	1.000
178529 | RC	s__Eubacterium rectale	1.67762E-05	2.91E-05	**< 0.0001**	1	**< 0.0001**
178529 | TI	s__Eubacterium rectale	0	0	1.000	**< 0.0001**	1
46159 | LC	s__Eubacterium rectale	0	0	1	**< 0.0001**	**< 0.0001**
46159 | RC	s__Eubacterium rectale	4.94473E-05	0.000113	**< 0.0001**	1	0.854
46159 | TI	s__Eubacterium rectale	3.77566E-05	8.66E-05	**< 0.0001**	0.854	1
177918 | LC	s__Faecalibacterium prausnitzii	0	0	1	**< 0.0001**	**< 0.0001**
172274 | LC	s__Faecalibacterium	0.000210881	0.000683	1	**< 0.0001**	0.677

	prausnitzii					
172274 | RC	s__Faecalibacterium prausnitzii	0	0	**< 0.0001**	1	**< 0.0001**
172274 | TI	s__Faecalibacterium prausnitzii	0.000342741	0.001097	0.677	**< 0.0001**	1
177918 | RC	s__Faecalibacterium prausnitzii	0.000366835	0.001585	**< 0.0001**	1	0.372
177918 | TI	s__Faecalibacterium prausnitzii	0.001305067	0.004368	**< 0.0001**	0.372	1
202045 | LC	s__Lactobacillus intestinalis	0	0	1	**< 0.0001**	1.000
202045 | RC	s__Lactobacillus intestinalis	1.70152E-05	2.27E-05	**< 0.0001**	1	**< 0.0001**
202045 | TI	s__Lactobacillus intestinalis	0	0	1.000	**< 0.0001**	1
470747 | LC	s__Oribacteriumsinus	0	0	1	1.000	**< 0.0001**
470747 | RC	s__Oribacteriumsinus	0	0	1.000	1	**< 0.0001**
470747 | TI	s__Oribacteriumsinus	8.89566E-05	0.000223	**< 0.0001**	**< 0.0001**	1
291090 | LC	s__Parabacteroides distasonis	0.002227145	0.007055	1	**< 0.0001**	0.605
291090 | RC	s__Parabacteroides distasonis	0	0	**< 0.0001**	1	**< 0.0001**
291090 | TI	s__Parabacteroides distasonis	0.000280852	0.000776	0.605	**< 0.0001**	1
160550 | LC	s__Photobacterium rosenbergii	0	0	1	1.000	**< 0.0001**
160550 | RC	s__Photobacterium rosenbergii	0	0	1.000	1	**< 0.0001**
160550 | TI	s__Photobacterium rosenbergii	0.000166292	0.000307	**< 0.0001**	**< 0.0001**	1
10167 | LC	s__Serratia entomophila	0.000167256	0.000522	1	**< 0.0001**	0.481
10167 | RC	s__Serratia entomophila	0	0	**< 0.0001**	1	**< 0.0001**
10167 | TI	s__Serratia entomophila	0.000738188	0.001888	0.481	**< 0.0001**	1
154802 | LC	s__Streptococcus pseudopneumoniae	0.000722807	0.001846	1	**0.014**	**0.013**
154802 | RC	s__Streptococcus pseudopneumoniae	0.002030324	0.00376	**0.014**	1	0.800
154802 | TI	s__Streptococcus pseudopneumoniae	0.003732517	0.010765	**0.013**	0.800	1

Appendix VIII: Left colon taxa that vary significantly in abundance in at least one of the three pairwise comparisons performed (controls vs. CD; controls vs. UC; and CD vs. UC). Kruskal-Wallis followed by Bonferroni correction for multiple comparisons test. Bonferroni corrected significance level: 0.0167. Values in bold indicate $p < 0.0167$. Observations indicate the number of samples from each group employed in the analysis. Mean and std. deviation indicate the mean and std. deviation of 16S reads obtained. *p*|CD, *p*|Control and *p*|UC indicate the *p* values obtained by comparison to the CD, control and UC, respectively.

Variable	Observ-ations	Mean	Std. deviation	*p*|CD	*p*| Control	*p*|UC
Phylum						
Actinobacteria | CD	39	0.018962	0.032603	1	0.294	0.081
Actinobacteria | Control	34	0.008434	0.010508	0.294	1	**0.014**
Actinobacteria | UC	13	0.025656	0.030998	0.081	**0.014**	1
Tenericutes | CD	39	0.02596	0.067082	1	**0.001**	0.710
Tenericutes | Control	34	0.034271	0.029421	**0.001**	1	0.040
Tenericutes | UC	13	0.05702	0.153666	0.710	0.040	1
Class						
Actinobacteria | CD	39	0.018962	0.032602	1	0.294	0.081
Actinobacteria | Control	34	0.008434	0.010508	0.294	1	**0.014**
Actinobacteria | UC	13	0.025656	0.030998	0.081	**0.014**	1
Erysipelotrichi | CD	39	0.020872	0.067177	1	**0.002**	0.519
Erysipelotrichi | Control	34	0.026331	0.030836	**0.002**	1	0.105
Erysipelotrichi | UC	13	0.051238	0.154714	0.519	0.105	1
Order						
Erysipelotrichales | CD	39	0.020872	0.067177	1	**0.002**	0.519
Erysipelotrichales | Control	34	0.026331	0.030836	**0.002**	1	0.105
Erysipelotrichales | UC	13	0.051238	0.154714	0.519	0.105	1
Gemellales | CD	39	0.007518	0.023821	1	**0.009**	0.708
Gemellales | Control	34	0.000169	0.000333	**0.009**	1	0.131
Gemellales | UC	13	0.002357	0.006995	0.708	0.131	1
Natranaerobiales | CD	39	4.55E-07	1.86E-06	1	0.203	0.059
Natranaerobiales | Control	34	0	0	0.203	1	**0.006**
Natranaerobiales | UC	13	1.57E-06	4.34E-06	0.059	**0.006**	1
Family						
Anaerobrancaceae | CD	39	4.55E-07	1.86E-06	1	0.203	0.059
Anaerobrancaceae | Control	34	0	0	0.203	1	**0.006**
Anaerobrancaceae | UC	13	1.57E-06	4.34E-06	0.059	**0.006**	1
Clostridiaceae | CD	39	0.013471	0.063212	1	**0.010**	0.768
Clostridiaceae | Control	34	0.015861	0.03601	**0.010**	1	0.032

Clostridiaceae | UC	13	0.005225	0.009242	0.768	0.032	1
Erysipelotrichaceae | CD	39	0.020872	0.067177	1	**0.002**	0.519
Erysipelotrichaceae | Control	34	0.026331	0.030836	**0.002**	1	0.105
Erysipelotrichaceae | UC	13	0.051238	0.154714	0.519	0.105	1
Gemellaceae | CD	39	0.007518	0.023821	1	**0.009**	0.708
Gemellaceae | Control	34	0.000169	0.000333	**0.009**	1	0.131
Gemellaceae | UC	13	0.002357	0.006995	0.708	0.131	1
Hyphomicrobiaceae | CD	39	0	0	1	0.497	**0.009**
Hyphomicrobiaceae | Control	34	1.24E-07	7.22E-07	0.497	1	0.039
Hyphomicrobiaceae | UC	13	1.83E-06	6.19E-06	**0.009**	0.039	1
Paenibacillaceae | CD	39	1.24E-05	3.32E-05	1	**0.008**	0.047
Paenibacillaceae | Control	34	3.56E-05	6.32E-05	**0.008**	1	0.973
Paenibacillaceae | UC	13	1.88E-05	1.86E-05	0.047	0.973	1
Veillonellaceae | CD	39	0.018433	0.032171	1	**0.004**	0.513
Veillonellaceae | Control	34	0.004089	0.007411	**0.004**	1	**0.006**
Veillonellaceae | UC	13	0.022864	0.034624	0.513	**0.006**	1
Genus						
A55_D21 | CD	39	4.55E-07	1.86E-06	1	0.203	0.059
A55_D21 | Control	34	0	0	0.203	1	**0.006**
A55_D21 | UC	13	1.57E-06	4.34E-06	0.059	**0.006**	1
Actinobacillus | CD	39	1.38E-05	4.41E-05	1	0.408	0.067
Actinobacillus | Control	34	6.42E-06	1.83E-05	0.408	1	**0.017**
Actinobacillus | UC	13	0.000244	0.000764	0.067	**0.017**	1
Alcanivorax | CD	39	0	0	1	1.000	**0.002**
Alcanivorax | Control	34	0	0	1.000	1	**0.002**
Alcanivorax | UC	13	1.54E-06	5.37E-06	**0.002**	**0.002**	1
Bradyrhizobium | CD	39	2.72E-08	1.7E-07	1	0.577	**0.014**
Bradyrhizobium | Control	34	1.73E-07	8.74E-07	0.577	1	0.044
Bradyrhizobium | UC	13	2.21E-06	7.6E-06	**0.014**	0.044	1
Clostridium | CD	39	0.013408	0.063224	1	0.021	0.343
Clostridium | Control	34	0.015727	0.035732	0.021	1	**0.009**
Clostridium | UC	13	0.003793	0.006415	0.343	**0.009**	1
Clostridium | CD	39	0.001999	0.005261	1	**0.005**	0.805
Clostridium | Control	34	0.003219	0.011092	**0.005**	1	0.023
Clostridium | UC	13	0.00075	0.001188	0.805	0.023	1
Congregibacter | CD	39	0	0	1	1.000	**0.002**
Congregibacter | Control	34	0	0	1.000	1	**0.002**
Congregibacter | UC	13	1.23E-07	3.14E-07	**0.002**	**0.002**	1
Faecalibacterium | CD	39	0.073973	0.072259	1	**0.007**	0.349
Faecalibacterium | Control	34	0.028111	0.045738	**0.007**	1	0.308
Faecalibacterium | UC	13	0.048139	0.053287	0.349	0.308	1
Fusibacter | CD	39	8.15E-08	3.75E-07	1	0.711	0.025
Fusibacter | Control	34	3.44E-08	2.01E-07	0.711	1	**0.013**

Fusibacter | UC	13	4.65E-07	9.34E-07	0.025	**0.013**	1
Gemella | CD	39	0.007518	0.023821	1	**0.009**	0.708
Gemella | Control	34	0.000169	0.000333	**0.009**	1	0.131
Gemella | UC	13	0.002357	0.006995	0.708	0.131	1
Haemophilus | CD	39	0.000158	0.000605	1	**0.011**	0.507
Haemophilus | Control	34	1.31E-05	3.16E-05	**0.011**	1	**0.013**
Haemophilus | UC	13	5.8E-05	0.000114	0.507	**0.013**	1
Hydrogenophaga | CD	39	0	0	1	**0.013**	0.378
Hydrogenophaga | Control	34	1.87E-07	4.84E-07	**0.013**	1	0.361
Hydrogenophaga | UC	13	4.84E-08	1.74E-07	0.378	0.361	1
Kingella | CD	39	1.35E-07	3.98E-07	1	0.762	**0.004**
Kingella | Control	34	2.66E-07	7.38E-07	0.762	1	**0.010**
Kingella | UC	13	1.29E-05	3.97E-05	**0.004**	**0.010**	1
Lachnospira | CD	39	0.001315	0.004088	1	**0.008**	0.325
Lachnospira | Control	34	0.003706	0.006813	**0.008**	1	0.340
Lachnospira | UC	13	0.005425	0.009609	0.325	0.340	1
Leptotrichia | CD	39	4.64E-05	0.000144	1	0.624	**0.003**
Leptotrichia | Control	34	4.16E-05	0.000188	0.624	1	**0.011**
Leptotrichia | UC	13	0.001526	0.004705	**0.003**	**0.011**	1
Nitrincola | CD	39	0.000252	0.001465	1	**0.008**	0.651
Nitrincola | Control	34	0.00014	0.000725	**0.008**	1	0.019
Nitrincola | UC	13	0.000878	0.002859	0.651	0.019	1
Paenibacillus | CD	39	1.22E-05	3.32E-05	1	**0.007**	0.052
Paenibacillus | Control	34	3.53E-05	6.31E-05	**0.007**	1	0.975
Paenibacillus | UC	13	1.82E-05	1.79E-05	0.052	0.975	1
Paracoccus | CD	39	0	0	1	0.507	**0.008**
Paracoccus | Control	34	2.48E-08	1.44E-07	0.507	1	0.035
Paracoccus | UC	13	1.04E-06	3.37E-06	**0.008**	0.035	1
Turicibacter | CD	39	5.39E-05	0.000325	1	**0.009**	0.753
Turicibacter | Control	34	8.28E-05	0.000238	**0.009**	1	0.116
Turicibacter | UC	13	9.83E-07	1.62E-06	0.753	0.116	1
Veillonella | CD	39	0.006041	0.013034	1	0.091	0.024
Veillonella | Control	34	0.000612	0.001132	0.091	1	**0.001**
Veillonella | UC	13	0.020966	0.03504	0.024	**0.001**	1
Vitreoscilla | CD	39	6.29E-06	2.27E-05	1	0.343	0.044
Vitreoscilla | Control	34	1.11E-05	4.44E-05	0.343	1	**0.008**
Vitreoscilla | UC	13	0.000364	0.001236	0.044	**0.008**	1

Appendix IX: Left colon core OTUs (0.75) that vary significantly in abundance in at least one of the three pairwise comparisons performed (controls vs. CD; controls vs. UC; and CD vs. UC). Kruskal-Wallis followed by Bonferroni correction for multiple comparisons test. Bonferroni corrected significance level: 0.0167. Values in bold indicate $p<$ 0.0167. Observations indicate the numer of samples from each group employed in the analysis. Mean and std. deviation indicate the mean and std. deviation of the 16S reads obtained. *p*|CD, *p*|Control and *p*|UC indicate the *p* values obtained by comparison to the CD, control and UC, respectively.

OTU|Variable	Taxonomy	Observ-ations	Mean	Std. deviation	*p*|CD	*p*| Control	*p*|UC
	Order						
196757 | CD	Bacteroidales	39	0.01550	0.04196	1	**0.013**	0.149
196757 | Control	Bacteroidales	34	0.02479	0.07867	**0.013**	1	0.711
196757 | UC	Bacteroidales	13	0.03406	0.08694	0.149	0.711	1
204932 | CD	Clostridiales	39	0.00067	0.00271	1	**0.007**	0.202
204932 | Control	Clostridiales	34	0.00331	0.00686	**0.007**	1	0.495
204932 | UC	Clostridiales	13	0.00371	0.00624	0.202	0.495	1
	Family						
177359 | CD	Lachnospiraceae	39	0.00012	0.00026	1	0.059	0.291
177359 | Control	Lachnospiraceae	34	0.00014	0.00027	0.059	1	**0.017**
177359 | UC	Lachnospiraceae	13	0.00003	0.00004	0.291	**0.017**	1
183879 | CD	Lachnospiraceae	39	0.00058	0.00127	1	0.093	0.189
183879 | Control	Lachnospiraceae	34	0.00169	0.00344	0.093	1	**0.012**
183879 | UC	Lachnospiraceae	13	0.00304	0.01026	0.189	**0.012**	1
186717 | CD	Lachnospiraceae	39	0.00137	0.00513	1	**0.012**	0.863
186717 | Control	Lachnospiraceae	34	0.00099	0.00122	**0.012**	1	0.049
186717 | UC	Lachnospiraceae	13	0.00063	0.00148	0.863	0.049	1
195336 | CD	Lachnospiraceae	39	0.00009	0.00034	1	**0.007**	0.726
195336 | Control	Lachnospiraceae	34	0.00079	0.00283	**0.007**	1	0.109
195336 | UC	Lachnospiraceae	13	0.00004	0.00007	0.726	0.109	1
204129 | CD	Lachnospiraceae	39	0.00014	0.00057	1	**0.005**	0.125
204129 | Control	Lachnospiraceae	34	0.00029	0.00058	**0.005**	1	0.605
204129 | UC	Lachnospiraceae	13	0.00063	0.00139	0.125	0.605	1
207390 | CD	Lachnospiraceae	39	0.00053	0.00086	1	**0.013**	0.635
207390 | Control	Lachnospiraceae	34	0.00214	0.00668	**0.013**	1	0.024
207390 | UC	Lachnospiraceae	13	0.00069	0.00146	0.635	0.024	1
212503 | CD	Lachnospiraceae	39	0.00038	0.00113	1	**0.015**	0.108
212503 | Control	Lachnospiraceae	34	0.00068	0.00148	**0.015**	1	0.871
212503 | UC	Lachnospiraceae	13	0.00066	0.00109	0.108	0.871	1
214980 | CD	Lachnospiraceae	39	0.00053	0.00156	1	**0.007**	0.201
214980 | Control	Lachnospiraceae	34	0.00171	0.00273	**0.007**	1	0.491
214980 | UC	Lachnospiraceae	13	0.00505	0.01496	0.201	0.491	1
293896 | CD	Lachnospiraceae	39	0.00014	0.00039	1	0.103	0.109
293896 | Control	Lachnospiraceae	34	0.00013	0.00039	0.103	1	**0.006**

293896 | UC	Lachnospiraceae	13	0.00000	0.00001	0.109	**0.006**	1
104026 | CD	Neisseriaceae	39	0.00001	0.00003	1	0.948	**0.007**
104026 | Control	Neisseriaceae	34	0.00000	0.00001	0.948	1	**0.007**
104026 | UC	Neisseriaceae	13	0.00002	0.00004	**0.007**	**0.007**	1
	Genus						
173965 | CD	Adlercreutzia	39	0.00003	0.00014	1	**0.004**	0.504
173965 | Control	Adlercreutzia	34	0.00005	0.00012	**0.004**	1	**0.006**
173965 | UC	Adlercreutzia	13	0.00000	0.00001	0.504	**0.006**	1
138006 | CD	Alistipes	39	0.00181	0.00422	1	0.528	**0.017**
138006 | Control	Alistipes	34	0.00327	0.00767	0.528	1	0.057
138006 | UC	Alistipes	13	0.00032	0.00096	**0.017**	0.057	1
171559 | CD	Bacteroides	39	0.00011	0.00041	1	**0.000**	0.211
171559 | Control	Bacteroides	34	0.00033	0.00085	**0.000**	1	0.123
171559 | UC	Bacteroides	13	0.00016	0.00046	0.211	0.123	1
268332 | CD	Bacteroides	39	0.00001	0.00003	1	**0.000**	0.022
268332 | Control	Bacteroides	34	0.00004	0.00008	**0.000**	1	0.767
268332 | UC	Bacteroides	13	0.00003	0.00003	0.022	0.767	1
331820 | CD	Bacteroides	39	0.02308	0.05435	1	**0.002**	0.219
331820 | Control	Bacteroides	34	0.07823	0.12782	**0.002**	1	0.300
331820 | UC	Bacteroides	13	0.09346	0.15995	0.219	0.300	1
181036 | CD	Blautia	39	0.00049	0.00094	1	0.072	0.195
181036 | Control	Blautia	34	0.00074	0.00121	0.072	1	**0.010**
181036 | UC	Blautia	13	0.00032	0.00105	0.195	**0.010**	1
184926 | CD	Blautia	39	0.00008	0.00019	1	**0.016**	0.263
184926 | Control	Blautia	34	0.00018	0.00026	**0.016**	1	0.531
184926 | UC	Blautia	13	0.00023	0.00044	0.263	0.531	1
206611 | CD	Blautia	39	0.00002	0.00006	1	**0.007**	0.576
206611 | Control	Blautia	34	0.00050	0.00125	**0.007**	1	**0.013**
206611 | UC	Blautia	13	0.00009	0.00027	0.576	**0.013**	1
291742 | CD	Blautia	39	0.00002	0.00003	1	**0.008**	0.738
291742 | Control	Blautia	34	0.00005	0.00010	**0.008**	1	0.024
291742 | UC	Blautia	13	0.00011	0.00035	0.738	0.024	1
295257 | CD	Blautia	39	0.00005	0.00010	1	0.059	0.247
295257 | Control	Blautia	34	0.00010	0.00016	0.059	1	**0.013**
295257 | UC	Blautia	13	0.00022	0.00071	0.247	**0.013**	1
363767 | CD	Blautia	39	0.00071	0.00442	1	0.040	0.062
363767 | Control	Blautia	34	0.00004	0.00013	0.040	1	**0.001**
363767 | UC	Blautia	13	0.00000	0.00000	0.062	**0.001**	1
182190 | CD	Clostridium	39	0.00031	0.00105	1	**0.009**	0.509
182190 | Control	Clostridium	34	0.00123	0.00482	**0.009**	1	0.217
182190 | UC	Clostridium	13	0.00103	0.00292	0.509	0.217	1
182205 | CD	Clostridium	39	0.00004	0.00009	1	**0.003**	0.599
182205 | Control	Clostridium	34	0.00020	0.00035	**0.003**	1	0.111
182205 | UC	Clostridium	13	0.00005	0.00010	0.599	0.111	1
190872 | CD	Clostridium	39	0.00261	0.00536	1	**0.002**	0.289
190872 | Control	Clostridium	34	0.01630	0.02558	**0.002**	1	0.227

190872 | UC	Clostridium	13	0.00514	0.00950	0.289	0.227	1
191687 | CD	Clostridium	39	0.00006	0.00016	1	**0.013**	0.928
191687 | Control	Clostridium	34	0.00017	0.00046	**0.013**	1	0.089
191687 | UC	Clostridium	13	0.00021	0.00073	0.928	0.089	1
193509 | CD	Clostridium	39	0.00224	0.00940	1	0.162	0.120
193509 | Control	Clostridium	34	0.00066	0.00177	0.162	1	**0.011**
193509 | UC	Clostridium	13	0.00003	0.00007	0.120	**0.011**	1
204180 | CD	Clostridium	39	0.00324	0.01172	1	**< 0.0001**	0.681
204180 | Control	Clostridium	34	0.00969	0.01631	**< 0.0001**	1	**0.001**
204180 | UC	Clostridium	13	0.00226	0.00522	0.681	**0.001**	1
211903 | CD	Clostridium	39	0.00203	0.00563	1	**0.008**	0.748
211903 | Control	Clostridium	34	0.00337	0.00596	**0.008**	1	0.111
211903 | UC	Clostridium	13	0.00160	0.00262	0.748	0.111	1
291561 | CD	Clostridium	39	0.00078	0.00262	1	0.182	**0.010**
291561 | Control	Clostridium	34	0.00017	0.00052	0.182	1	0.113
291561 | UC	Clostridium	13	0.00473	0.01029	**0.010**	0.113	1
294710 | CD	Clostridium	39	0.00003	0.00013	1	**0.013**	0.583
294710 | Control	Clostridium	34	0.00008	0.00021	**0.013**	1	0.210
294710 | UC	Clostridium	13	0.00007	0.00023	0.583	0.210	1
299668 | CD	Clostridium	39	0.00020	0.00113	1	**0.006**	0.576
299668 | Control	Clostridium	34	0.00256	0.00774	**0.006**	1	**0.012**
299668 | UC	Clostridium	13	0.00003	0.00008	0.576	**0.012**	1
178146 | CD	Coprococcus	39	0.00001	0.00002	1	**0.001**	0.503
178146 | Control	Coprococcus	34	0.00003	0.00009	**0.001**	1	0.082
178146 | UC	Coprococcus	13	0.00002	0.00005	0.503	0.082	1
185703 | CD	Coprococcus	39	0.00003	0.00009	1	**0.013**	0.280
185703 | Control	Coprococcus	34	0.00010	0.00023	**0.013**	1	**0.004**
185703 | UC	Coprococcus	13	0.00001	0.00002	0.280	**0.004**	1
195196 | CD	Coprococcus	39	0.00012	0.00030	1	**0.005**	0.357
195196 | Control	Coprococcus	34	0.00036	0.00056	**0.005**	1	0.263
195196 | UC	Coprococcus	13	0.00018	0.00026	0.357	0.263	1
175261 | CD	Eubacterium	39	0.00189	0.00404	1	0.034	0.254
175261 | Control	Eubacterium	34	0.00222	0.00295	0.034	1	**0.008**
175261 | UC	Eubacterium	13	0.00071	0.00142	0.254	**0.008**	1
307868 | CD	Eubacterium	39	0.00007	0.00016	1	**0.015**	0.569
307868 | Control	Eubacterium	34	0.00011	0.00018	**0.015**	1	0.232
307868 | UC	Eubacterium	13	0.00011	0.00028	0.569	0.232	1
64396 | CD	Fusobacterium	39	0.01735	0.07009	1	**0.004**	0.913
64396 | Control	Fusobacterium	34	0.00229	0.01180	**0.004**	1	0.029
64396 | UC	Fusobacterium	13	0.00240	0.00568	0.913	0.029	1
470462 | CD	Gemella	39	0.00722	0.02350	1	**0.012**	0.649
470462 | Control	Gemella	34	0.00016	0.00032	**0.012**	1	0.173
470462 | UC	Gemella	13	0.00213	0.00652	0.649	0.173	1
74391 | CD	Leptotrichia	39	0.00003	0.00014	1	0.369	**0.000**

74391 | Control	Leptotrichia	34	0.00001	0.00003	0.369	1	**0.003**
74391 | UC	Leptotrichia	13	0.00145	0.00465	**0.000**	**0.003**	1
102362 | CD	Nitrincola	39	0.00025	0.00147	1	**0.009**	0.612
102362 | Control	Nitrincola	34	0.00014	0.00072	**0.009**	1	0.017
102362 | UC	Nitrincola	13	0.00088	0.00286	0.612	0.017	1
145149 | CD	Veillonella	39	0.00370	0.00873	1	0.067	0.047
145149 | Control	Veillonella	34	0.00027	0.00069	0.067	1	**0.001**
145149 | UC	Veillonella	13	0.00711	0.01031	0.047	**0.001**	1
585419 | CD	Veillonella	39	0.00234	0.00601	1	0.148	**0.011**
585419 | Control	Veillonella	34	0.00034	0.00060	0.148	1	**0.000**
585419 | UC	Veillonella	13	0.01382	0.02828	**0.011**	**0.000**	1
	Species						
344553 | CD	Clostridium hiranonis	39	0.00001	0.00003	1	**0.008**	0.101
344553 | Control	Clostridium hiranonis	34	0.00003	0.00007	**0.008**	1	0.776
344553 | UC	Clostridium hiranonis	13	0.00001	0.00001	0.101	0.776	1
15716 | CD	Clostridium spiroforme	39	0.00144	0.00778	1	0.072	0.103
15716 | Control	Clostridium spiroforme	34	0.00128	0.00315	0.072	1	**0.004**
15716 | UC	Clostridium spiroforme	13	0.00002	0.00005	0.103	**0.004**	1
109018 | CD	Eubacterium rectale	39	0.00287	0.00930	1	**0.003**	0.698
109018 | Control	Eubacterium rectale	34	0.00667	0.01735	**0.003**	1	0.080
109018 | UC	Eubacterium rectale	13	0.00603	0.01859	0.698	0.080	1
175682 | CD	Eubacterium rectale	39	0.00192	0.00951	1	**0.006**	0.581
175682 | Control	Eubacterium rectale	34	0.00314	0.00863	**0.006**	1	0.148
175682 | UC	Eubacterium rectale	13	0.00220	0.00541	0.581	0.148	1
181379 | CD	Eubacterium rectale	39	0.00213	0.00971	1	**0.008**	0.426
181379 | Control	Eubacterium rectale	34	0.00043	0.00127	**0.008**	1	**0.007**
181379 | UC	Eubacterium rectale	13	0.00010	0.00032	0.426	**0.007**	1
182994 | CD	Eubacterium rectale	39	0.00124	0.00284	1	**0.003**	0.242
182994 | Control	Eubacterium rectale	34	0.00850	0.01577	**0.003**	1	0.331
182994 | UC	Eubacterium rectale	13	0.00347	0.00737	0.242	0.331	1
198176 | CD	Eubacterium rectale	39	0.00244	0.00560	1	0.128	0.078
198176 | Control	Eubacterium rectale	34	0.00195	0.00269	0.128	1	**0.005**
198176 | UC	Eubacterium rectale	13	0.00102	0.00305	0.078	**0.005**	1
109807 | CD	MLG480	39	0.00019	0.00063	1	**0.012**	0.742
109807 | Control	MLG481	34	0.00102	0.00273	**0.012**	1	0.139
109807 | UC	MLG482	13	0.00047	0.00142	0.742	0.139	1

Appendix X: Left colon taxa that vary significantly in abundance in CD patients in at least one of the three pairwise comparisons performed (mild vs. moderate; mild vs. severe; and moderate vs. severe). Kruskal-Wallis followed by Bonferroni correction for multiple comparisons test. Bonferroni corrected significance level: 0.0167. Values in bold indicate $p < 0.0167$. Observations indicate the number of samples from each group employed in the analysis. Mean and std. deviation indicate the mean and std. deviation of the 16S reads obtained. *p*| Mild CD, *p*| Moderate CD and *p*| Severe CD indicate the *p* values obtained by comparison to the mild CD, moderate CD and severe CD, respectively.

Taxa|Variable	Observation	Mean	Std. deviation	*p*| Mild CD	*p*| Moderate CD	*p*| Severe CD
Class						
Alphaproteobacteria | mild	10	0.00000100	0.00000230	1	**0.011**	0.289
Alphaproteobacteria | moderate	6	0.00011129	0.00019806	**0.011**	1	0.047
Alphaproteobacteria | Severe	23	0.00001524	0.00004906	0.289	0.047	1
Clostridia | mild	10	0.01621754	0.01253036	1	0.799	**0.014**
Clostridia | moderate	6	0.01320299	0.00935826	0.799	1	0.081
Clostridia | Severe	23	0.00541300	0.00860916	**0.014**	0.081	1
Family						
ClostridialesFamilyXI.IncertaeSedis | mild	10	0.00000806	0.00000751	1	0.052	**0.005**
ClostridialesFamilyXI.IncertaeSedis | moderate	6	0.00011105	0.00015256	0.052	1	0.888
ClostridialesFamilyXI.IncertaeSedis | Severe	23	0.00124370	0.00333471	**0.005**	0.888	1
ClostridialesFamilyXIII.IncertaeSedis | mild	10	0.00000114	0.00000192	1	**0.010**	0.053
ClostridialesFamilyXIII.IncertaeSedis | moderate	6	0.00004458	0.00005044	**0.010**	1	0.190
ClostridialesFamilyXIII.IncertaeSedis | Severe	23	0.00001549	0.00005094	0.053	0.190	1
Hydrogenophilaceae | mild	10	0.00000016	0.00000049	1	0.115	0.304
Hydrogenophilaceae | moderate	6	0.00000026	0.00000041	0.115	1	**0.009**
Hydrogenophilaceae | Severe	23	0.00000000	0.00000000	0.304	**0.009**	1
Paenibacillaceae | mild	10	0.00000000	0.00000000	1	**0.004**	1.000
Paenibacillaceae | moderate	6	0.00000096	0.00000204	**0.004**	1	**0.001**
Paenibacillaceae | Severe	23	0.00000000	0.00000000	1.000	**0.001**	1
Genus						
Allobaculum | mild	10	0.00000010	0.00000031	1	0.846	**0.006**
Allobaculum | moderate	6	0.00000019	0.00000046	0.846	1	0.041
Allobaculum | Severe	23	0.00000452	0.00000824	**0.006**	0.041	1
Atopobium | mild	10	0.00016881	0.00035816	1	0.105	0.200
Atopobium | moderate	6	0.00000839	0.00001398	0.105	1	**0.004**

Atopobium | Severe	23	0.00015067	0.00031435	0.200	**0.004**	1
Dialister | mild	10	0.00075767	0.00113885	1	**0.014**	0.277
Dialister | moderate	6	0.01321420	0.02444939	**0.014**	1	0.061
Dialister | Severe	23	0.01179565	0.02585354	0.277	0.061	1
Facklamia | mild	10	0.00014003	0.00028909	1	0.153	0.233
Facklamia | moderate	6	0.00010991	0.00016564	0.153	1	**0.009**
Facklamia | Severe	23	0.00001048	0.00004513	0.233	**0.009**	1
Hydrogenophilus | mild	10	0.00000015	0.00000049	1	0.115	0.304
Hydrogenophilus | moderate	6	0.00000026	0.00000041	0.115	1	**0.009**
Hydrogenophilus | Severe	23	0.00000000	0.00000000	0.304	**0.009**	1
Pantoea | mild	10	0.00000011	0.00000033	1	0.542	**0.006**
Pantoea | moderate	6	0.00000036	0.00000057	0.542	1	0.116
Pantoea | Severe	23	0.00000329	0.00000568	**0.006**	0.116	1
Species						
Allobaculum spID4 | mild	10	0.00000010	0.00000031	1	0.846	**0.006**
Allobaculum spID4 | moderate	6	0.00000019	0.00000046	0.846	1	0.041
Allobaculums pID4 | Severe	23	0.00000452	0.00000824	**0.006**	0.041	1
Atopobium parvulum | mild	10	0.00008686	0.00016267	1	0.116	0.213
Atopobium parvulum | moderate	6	0.00000662	0.00001155	0.116	1	**0.005**
Atopobium parvulum | Severe	23	0.00008780	0.00019998	0.213	**0.005**	1
Hydrogenophilus thermoluteolus | mild	10	0.00000015	0.00000049	1	0.115	0.304
Hydrogenophilus thermoluteolus | moderate	6	0.00000026	0.00000041	0.115	1	**0.009**
Hydrogenophilus thermoluteolus | Severe	23	0.00000000	0.00000000	0.304	**0.009**	1
Oribacterium sinus | mild	10	0.00000148	0.00000223	1	0.841	**0.009**
Oribacterium sinus | moderate	6	0.00000213	0.00000237	0.841	1	0.053
Oribacterium sinus | Severe	23	0.00000912	0.00001533	**0.009**	0.053	1
Selenomonas sputigena | mild	10	0.00000000	0.00000000	1	**0.005**	0.488
Selenomonas sputigena | moderate	6	0.00000051	0.00000060	**0.005**	1	**0.010**
Selenomonas sputigena | Severe	23	0.00000083	0.00000377	0.488	**0.010**	1
Streptococcus pneumoniae | mild	10	0.00051121	0.00118199	1	0.019	0.790
Streptococcus pneumoniae | moderate	6	0.00001163	0.00000812	0.019	1	**0.015**
Streptococcus pneumoniae | Severe	23	0.00043005	0.00149212	0.790	**0.015**	1

Appendix XI: Left colon taxa that vary significantly in abundance in UC patients in at least one of the three pairwise comparisons performed (mild vs. moderate; mild vs. severe; and moderate vs. severe). Kruskal-Wallis followed by Bonferroni correction for multiple comparisons test. Bonferroni corrected significance level: 0.0167. Values in bold indicate $p< 0.0167$. Observations indicate the number of samples from each group employed in the analysis. Mean and std. deviation indicate the mean and std. deviation of the 16S reads obtained. *p*| Mild UC, *p*| Moderate UC and *p*| Severe UC indicate the *p* values obtained by comparison to the mild UC, moderate UC and severe UC, respectively.

Variable	Observations	Mean	Std. deviation	*p*| Mild UC	*p*| Moderate UC	*p*| Severe UC
Phylum						
Fibrobacteres | mild	4	0.00000000	0.00000000	1	1.000	**0.017**
Fibrobacteres | moderate	5	0.00000000	0.00000000	1.000	1	**0.012**
Fibrobacteres | Severe	4	0.00000364	0.00000483	**0.017**	**0.012**	1
Class						
Bacteroidia | mild	4	0.12571741	0.12458179	1	0.056	**0.014**
Bacteroidia | moderate	5	0.01292891	0.01636394	0.056	1	0.503
Bacteroidia | Severe	4	0.00116030	0.00038127	**0.014**	0.503	1
Fibrobacteres | mild	4	0.00000000	0.00000000	1	1.000	**0.017**
Fibrobacteres | moderate	5	0.00000000	0.00000000	1.000	1	**0.012**
Fibrobacteres | Severe	4	0.00000364	0.00000483	**0.017**	**0.012**	1
Family						
Enterococcaceae | mild	4	0.00000080	0.00000160	1	0.073	**0.003**
Enterococcaceae | moderate	5	0.00023780	0.00049621	0.073	1	0.197
Enterococcaceae | Severe	4	0.00090038	0.00094107	**0.003**	0.197	1
Fibrobacteraceae | mild	4	0.00000000	0.00000000	1	1.000	**0.017**
Fibrobacteraceae | moderate	5	0.00000000	0.00000000	1.000	1	**0.012**
Fibrobacteraceae | Severe	4	0.00000364	0.00000483	**0.017**	**0.012**	1
Lachnospiraceae | mild	4	0.07654722	0.03813620	1	0.180	**0.011**
Lachnospiraceae | moderate	5	0.03843633	0.04405575	0.180	1	0.180
Lachnospiraceae | Severe	4	0.00468567	0.00480477	**0.011**	0.180	1
Neisseriaceae | mild	4	0.00000867	0.00000701	1	**0.008**	0.029
Neisseriaceae | moderate	5	0.00176812	0.00330829	**0.008**	1	0.730
Neisseriaceae | Severe	4	0.03582059	0.07145212	0.029	0.730	1
Genus						
Abiotrophia | mild	4	0.00000159	0.00000122	1	0.096	**0.014**
Abiotrophia | moderate	5	0.00007269	0.00012829	0.096	1	0.358
Abiotrophia | Severe	4	0.00015778	0.00021169	**0.014**	0.358	1
Carnobacterium | mild	4	0.00000084	0.00000168	1	0.197	**0.005**
Carnobacterium | moderate	5	0.00001395	0.00002043	0.197	1	0.095
Carnobacterium | Severe	4	0.00009431	0.00005986	**0.005**	0.095	1
Coprococcus | mild	4	0.03214956	0.02602387	1	**0.016**	**0.014**

Coprococcus | moderate	5	0.00158831	0.00095911	**0.016**	1	0.863
Coprococcus | Severe	4	0.00254711	0.00337017	**0.014**	0.863	1
Enterococcus | mild	4	0.00000020	0.00000040	1	0.087	**0.003**
Enterococcus | moderate	5	0.00002988	0.00003928	0.087	1	0.144
Enterococcus | Severe	4	0.00027455	0.00024080	**0.003**	0.144	1
Eubacterium | mild	4	0.00006710	0.00011305	1	**0.012**	**0.017**
Eubacterium | mild	4	0.01325109	0.01262310	1	0.243	**0.014**
Eubacterium | moderate	5	0.00000000	0.00000000	**0.012**	1	1.000
Eubacterium | moderate	5	0.00202095	0.00164487	0.243	1	0.157
Eubacterium | Severe	4	0.00000000	0.00000000	**0.017**	1.000	1
Eubacterium | Severe	4	0.00057371	0.00105814	**0.014**	0.157	1
Fibrobacter | mild	4	0.00000000	0.00000000	1	1.000	**0.017**
Fibrobacter | moderate	5	0.00000000	0.00000000	1.000	1	**0.012**
Fibrobacter | Severe	4	0.00000364	0.00000483	**0.017**	**0.012**	1
Lachnospira | mild	4	0.00678942	0.01023953	1	1.000	0.018
Lachnospira | moderate	5	0.00890624	0.01221659	1.000	1	**0.013**
Lachnospira | Severe	4	0.00000547	0.00000715	0.018	**0.013**	1
Macrococcus | mild	4	0.00000020	0.00000040	1	0.559	0.076
Macrococcus | moderate	5	0.00000000	0.00000000	0.559	1	**0.014**
Macrococcus | Severe	4	0.00000674	0.00000999	0.076	**0.014**	1
Neisseria | mild	4	0.00000439	0.00000301	1	**0.008**	0.029
Neisseria | moderate	5	0.00164043	0.00321068	**0.008**	1	0.730
Neisseria | Severe	4	0.00876659	0.01747162	0.029	0.730	1
Vagococcus | mild	4	0.00000000	0.00000000	1	0.033	**0.004**
Vagococcus | moderate	5	0.00017569	0.00038772	0.033	1	0.382
Vagococcus | Severe	4	0.00020743	0.00034307	**0.004**	0.382	1
Species						
Abiotrophia defectiva | mild	4	0.00000159	0.00000122	1	0.096	**0.014**
Abiotrophia defectiva | moderate	5	0.00007269	0.00012829	0.096	1	0.358
Abiotrophia defectiva | Severe	4	0.00015778	0.00021169	**0.014**	0.358	1
Clostridium hylemonae | mild	4	0.00002938	0.00004356	1	**0.012**	**0.017**
Clostridium hylemonae | moderate	5	0.00000000	0.00000000	**0.012**	1	1.000
Clostridium hylemonae | Severe	4	0.00000000	0.00000000	**0.017**	1.000	1
Corynebacterium durum | mild	4	0.00000000	0.00000000	1	**0.015**	0.050
Corynebacterium durum | moderate	5	0.00002260	0.00003579	**0.015**	1	0.708
Corynebacterium durum | Severe	4	0.00002240	0.00003212	0.050	0.708	1
Dialister pneumosintes | mild	4	0.00009441	0.00017132	1	**0.005**	0.044
Dialister pneumosintes | moderate	5	0.00000000	0.00000000	**0.005**	1	0.478
Dialister pneumosintes | Severe	4	0.00000270	0.00000539	0.044	0.478	1
Enterococcus haemoperoxidus | mild	4	0.00000060	0.00000120	1	0.487	**0.015**
Enterococcus haemoperoxidus | moderate	5	0.00003191	0.00006944	0.487	1	0.061
Enterococcus haemoperoxidus | Severe	4	0.00041739	0.00078945	**0.015**	0.061	1

Fibrobacter intestinalis | mild	4	0.00000000	0.00000000	1	1.000	**0.017**
Fibrobacter intestinalis | moderate	5	0.00000000	0.00000000	1.000	1	**0.012**
Fibrobacter intestinalis | Severe	4	0.00000364	0.00000483	**0.017**	**0.012**	1
Macrococcus caseolyticus | mild	4	0.00000020	0.00000040	1	0.559	0.076
Macrococcus caseolyticus | moderate	5	0.00000000	0.00000000	0.559	1	**0.014**
Macrococcus caseolyticus | Severe	4	0.00000674	0.00000999	0.076	**0.014**	1
MLG480 | mild	4	0.00024802	0.00044472	1	0.128	0.361
MLG480 | moderate	5	0.00000076	0.00000106	0.128	1	**0.013**
MLG480 | Severe	4	0.00131218	0.00254768	0.361	**0.013**	1
Neisseria bacilliformis | mild	4	0.00000293	0.00000185	1	**0.007**	0.037
Neisseria bacilliformis | moderate	5	0.00163219	0.00319456	**0.007**	1	0.605
Neisseria bacilliformis | Severe	4	0.00775015	0.01544479	0.037	0.605	1
Staphylococcus aureus | mild	4	0.00000000	0.00000000	1	1.000	**0.004**
Staphylococcus aureus | moderate	5	0.00000000	0.00000000	1.000	1	**0.002**
Staphylococcus aureus | Severe	4	0.00000607	0.00000774	**0.004**	**0.002**	1

Appendix XII: Right colon taxa that vary significantly in abundance in at least one of the three pairwise comparisons performed (controls vs. CD; controls vs. UC; and CD vs. UC). Kruskal-Wallis followed by Bonferroni correction for multiple comparisons. Values in bold indicate $p <$ Bonferroni corrected significance level of 0.0167. Observations indicate the number of samples from each group employed in the analysis. Mean and std. deviation indicate the mean and std. deviation of the 16S reads obtained. *p*| Control, *p*| CD and *p*| UC indicate the *p* values obtained by comparison to the control, CD and UC, respectively.

Variable	Observations	Mean	Std. deviation	*p* | Control	*p* | UC	*p* | CD
PHYLUM						
Firmicutes | Control	21	0.600099	0.190174	1	**0.011**	0.246
Firmicutes | UC	14	0.408788	0.165937	**0.011**	1	0.075
Firmicutes | CD	37	0.518395	0.20456	0.246	0.075	1
CLASS						
Negativicutes | Control	21	0.003111	0.007177	1	0.020	**0.008**
Negativicutes | UC	14	0.021227	0.0287	0.020	1	0.799
Negativicutes | CD	37	0.017518	0.030741	**0.008**	0.799	1
Clostridia | Control	21	0.580535	0.184267	1	**0.006**	0.054
Clostridia | UC	14	0.373121	0.169775	**0.006**	1	0.173
Clostridia | CD	37	0.454472	0.218102	0.054	0.173	1
Verrucomicrobiae | Control	21	0.000171	0.000503	1	**0.014**	0.296
Verrucomicrobiae | UC	14	2.14E-06	6.9E-06	**0.014**	1	0.075
Verrucomicrobiae | CD	37	0.000192	0.000715	0.296	0.075	1
Betaproteobacteria | Control	21	0.008432	0.021154	1	0.120	**0.002**
Betaproteobacteria | UC	14	0.025282	0.046107	0.120	1	0.306
Betaproteobacteria | CD	37	0.028815	0.052705	**0.002**	0.306	1
		0	0			
ORDER		0	0			
Pasteurellales | Control	21	4.85E-05	0.000194	1	**0.003**	**0.002**
Pasteurellales | UC	14	0.001324	0.003365	**0.003**	1	0.544
Pasteurellales | CD	37	0.00015	0.00038	**0.002**	0.544	1
Chromatiales | Control	21	9.52E-08	4.36E-07	1	**0.003**	**0.009**
Chromatiales | UC	14	4.14E-06	6.81E-06	**0.003**	1	0.331
Chromatiales | CD	37	4.11E-06	1.13E-05	**0.009**	0.331	1
Burkholderiales | Control	21	0.0083	0.021182	1	0.764	**0.015**
Burkholderiales | UC	14	0.01366	0.029981	0.764	1	0.073
Burkholderiales | CD	37	0.027989	0.052729	**0.015**	0.073	1
Selenomonadales | Control	21	0.003111	0.007177	1	0.020	**0.008**
Selenomonadales | UC	14	0.021227	0.0287	0.020	1	0.799
Selenomonadales | CD	37	0.017518	0.030741	**0.008**	0.799	1
Clostridiales | Control	21	0.580535	0.184267	1	**0.006**	0.054

Clostridiales | UC	14	0.373118	0.169775	**0.006**	1	0.173
Clostridiales | CD	37	0.454468	0.218104	0.054	0.173	1
Hydrogenophilales | Control	21	0	0	1	**0.008**	0.660
Hydrogenophilales | UC	14	5.71E-07	1.22E-06	**0.008**	1	**0.011**
Hydrogenophilales | CD	37	1.08E-07	6.58E-07	0.660	**0.011**	1
Oceanospirillales | Control	21	0.000762	0.002181	1	**0.007**	0.127
Oceanospirillales | UC	14	0.002621	0.005193	**0.007**	1	0.100
Oceanospirillales | CD	37	0.000361	0.000828	0.127	0.100	1
Rhizobiales | Control	21	4.76E-07	1.08E-06	1	**0.001**	0.537
Rhizobiales | UC	14	4.76E-05	0.00011	**0.001**	1	**0.002**
Rhizobiales | CD	37	8.49E-06	2.85E-05	0.537	**0.002**	1
Verrucomicrobiales | Control	21	0.000171	0.000503	1	**0.014**	0.296
Verrucomicrobiales | UC	14	2.14E-06	6.9E-06	**0.014**	1	0.075
Verrucomicrobiales | CD	37	0.000192	0.000715	0.296	0.075	1
FAMILY						
Verrucomicrobiaceae | Control	21	0.000171	0.000503	1	**0.014**	0.296
Verrucomicrobiaceae | UC	14	2.14E-06	6.9E-06	**0.014**	1	0.075
Verrucomicrobiaceae | CD	37	0.000192	0.000715	0.296	0.075	1
Staphylococcaceae | Control	21	0.000711	0.00182	1	0.345	**0.009**
Staphylococcaceae | UC	14	0.000118	0.000161	0.345	1	0.216
Staphylococcaceae | CD	37	0.000445	0.001869	**0.009**	0.216	1
Lachnospiraceae | Control	21	0.147464	0.114625	1	**0.006**	0.058
Lachnospiraceae | UC	14	0.053922	0.051441	**0.006**	1	0.178
Lachnospiraceae | CD	37	0.094256	0.09088	0.058	0.178	1
Halomonadaceae | Control	21	5.33E-06	2.4E-05	1	**0.014**	0.263
Halomonadaceae | UC	14	0.000106	0.000379	**0.014**	1	0.083
Halomonadaceae | CD	37	1.62E-06	4.49E-06	0.263	0.083	1
Pasteurellaceae | Control	21	4.85E-05	0.000194	1	**0.003**	**0.002**
Pasteurellaceae | UC	14	0.001324	0.003365	**0.003**	1	0.544
Pasteurellaceae | CD	37	0.00015	0.00038	**0.002**	0.544	1
Paenibacillaceae | Control	21	0.000064	8.83E-05	1	**0.014**	**0.009**
Paenibacillaceae | UC	14	1.66E-05	2.89E-05	**0.014**	1	0.658
Paenibacillaceae | CD	37	4.75E-05	0.000202	**0.009**	0.658	1
Listeriaceae | Control	21	2.76E-06	6.02E-06	1	0.310	**0.011**
Listeriaceae | UC	14	5.71E-07	1.65E-06	0.310	1	0.272
Listeriaceae | CD	37	1.08E-07	6.58E-07	**0.011**	0.272	1
Bradyrhizobiaceae | Control	21	3.81E-07	1.02E-06	1	**0.006**	0.604
Bradyrhizobiaceae | UC	14	7.43E-06	1.78E-05	**0.006**	1	**0.010**

Bradyrhizobiaceae | CD	37	6.43E-06	2.36E-05	0.604	**0.010**	1
Methylococcaceae | Control	21	3.81E-07	1.02E-06	1	**0.004**	0.243
Methylococcaceae | UC	14	6.19E-05	0.000189	**0.004**	1	0.032
Methylococcaceae | CD	37	4.86E-07	1.66E-06	0.243	0.032	1
Hydrogenophilaceae | Control	21	0	0	1	**0.008**	0.660
Hydrogenophilaceae | UC	14	5.71E-07	1.22E-06	**0.008**	1	**0.011**
Hydrogenophilaceae | CD	37	1.08E-07	6.58E-07	0.660	**0.011**	1
GENUS						
Porphyromonas | Control	21	2.86E-07	7.17E-07	1	**0.002**	0.443
Porphyromonas | UC	14	1.26E-05	2.41E-05	**0.002**	1	**0.007**
Porphyromonas | CD	37	0.000002	5.93E-06	0.443	**0.007**	1
Lautropia | Control	21	9.52E-08	4.36E-07	1	**0.001**	0.555
Lautropia | UC	14	1.06E-05	3.02E-05	**0.001**	1	**< 0.0001**
Lautropia | CD	37	0	0	0.555	**< 0.0001**	1
Methylobacterium | Control	21	9.52E-08	4.36E-07	1	**0.004**	0.243
Methylobacterium | UC	14	2.56E-05	6.35E-05	**0.004**	1	0.032
Methylobacterium | CD	37	0.000002	5.52E-06	0.243	0.032	1
Akkermansia | Control	21	0.000171	0.000503	1	**0.014**	0.296
Akkermansia | UC	14	2.14E-06	6.9E-06	**0.014**	1	0.075
Akkermansia | CD	37	0.000192	0.000715	0.296	0.075	1
Tannerella | Control	21	5.71E-06	2.62E-05	1	0.042	0.431
Tannerella | UC	14	4.29E-07	8.52E-07	0.042	1	**0.004**
Tannerella | CD	37	0	0	0.431	**0.004**	1
Haemophilus | Control	21	2.76E-06	5.88E-06	1	**0.007**	**0.010**
Haemophilus | UC	14	0.000159	0.000382	**0.007**	1	0.478
Haemophilus | CD	37	6.11E-05	0.000203	**0.010**	0.478	1
Finegoldia | Control	21	9.52E-08	4.36E-07	1	**0.014**	0.941
Finegoldia | UC	14	6.13E-05	0.000169	**0.014**	1	**0.009**
Finegoldia | CD	37	1.08E-07	4.58E-07	0.941	**0.009**	1
Turicibacter | Control	21	0.00014	0.000397	1	**0.007**	**0.006**
Turicibacter | UC	14	5.71E-07	1.22E-06	**0.007**	1	0.569
Turicibacter | CD	37	7.98E-05	0.000479	**0.006**	0.569	1
Nitrincola | Control	21	1.35E-05	4.13E-05	1	**0.008**	0.021
Nitrincola | UC	14	0.000794	0.002318	**0.008**	1	0.360
Nitrincola | CD	37	7.74E-05	0.000351	0.021	0.360	1
Hydrogenophilus | Control	21	0	0	1	**0.008**	0.660
Hydrogenophilus | UC	14	5.71E-07	1.22E-06	**0.008**	1	**0.011**
Hydrogenophilus | CD	37	1.08E-07	6.58E-07	0.660	**0.011**	1
Listeria | Control	21	2.76E-06	6.02E-06	1	0.310	**0.011**
Listeria | UC	14	5.71E-07	1.65E-06	0.310	1	0.272

Listeria | CD	37	1.08E-07	6.58E-07	**0.011**	0.272	1
Actinobacillus | Control	21	1.14E-06	3.07E-06	1	**0.005**	0.018
Actinobacillus | UC	14	0.000278	0.000961	**0.005**	1	0.307
Actinobacillus | CD	37	3.13E-05	9.19E-05	0.018	0.307	1
Anaerococcus | Control	21	2.19E-06	8.72E-06	1	**0.001**	0.054
Anaerococcus | UC	14	0.003481	0.012625	**0.001**	1	0.065
Anaerococcus | CD	37	9.26E-05	0.000526	0.054	0.065	1
Catonella | Control	21	1.05E-06	3.5E-06	1	0.133	0.231
Catonella | UC	14	4.14E-06	7.98E-06	0.133	1	**0.007**
Catonella | CD	37	1.19E-06	6.3E-06	0.231	**0.007**	1
Mobiluncus | Control	21	0	0	1	**0.007**	0.677
Mobiluncus | UC	14	0.000001	2.18E-06	**0.007**	1	**0.009**
Mobiluncus | CD	37	5.41E-08	3.29E-07	0.677	**0.009**	1
Pantoea | Control	21	0	0	1	0.058	**0.002**
Pantoea | UC	14	2.14E-06	4.8E-06	0.058	1	0.523
Pantoea | CD	37	6.86E-06	3.24E-05	**0.002**	0.523	1
Enterobacter | Control	21	3.62E-06	7.66E-06	1	**0.003**	0.109
Enterobacter | UC	14	0.001558	0.005649	**0.003**	1	0.057
Enterobacter | CD	37	0.000212	0.001178	0.109	0.057	1
Paenibacillus | Control	21	6.39E-05	8.83E-05	1	**0.013**	**0.007**
Paenibacillus | UC	14	1.64E-05	2.88E-05	**0.013**	1	0.699
Paenibacillus | CD	37	4.72E-05	0.000202	**0.007**	0.699	1
Staphylococcus | Control	21	0.000708	0.001821	1	0.224	**0.010**
Staphylococcus | UC	14	0.000117	0.000162	0.224	1	0.370
Staphylococcus | CD	37	0.000429	0.001851	**0.010**	0.370	1
Vitreoscilla | Control	21	4.57E-06	1.61E-05	1	**0.009**	0.319
Vitreoscilla | UC	14	0.000257	0.000887	**0.009**	1	0.044
Vitreoscilla | CD	37	3.24E-06	6.38E-06	0.319	0.044	1
Alcanivorax | Control	21	0	0	1	**0.012**	1.000
Alcanivorax | UC	14	1.29E-06	3.29E-06	**0.012**	1	**0.006**
Alcanivorax | CD	37	0	0	1.000	**0.006**	1
Veillonella | Control	21	0.000375	0.000585	1	**0.003**	0.106
Veillonella | UC	14	0.017971	0.027592	**0.003**	1	0.063
Veillonella | CD	37	0.004214	0.00893	0.106	0.063	1
Tatumella | Control	21	6.76E-06	2.67E-05	1	**0.014**	0.090
Tatumella | UC	14	0.000351	0.001004	**0.014**	1	0.224
Tatumella | CD	37	1.01E-05	2.78E-05	0.090	0.224	1
Afipia | Control	21	1.9E-07	6.02E-07	1	0.027	0.913
Afipia | UC	14	4.86E-06	1.59E-05	0.027	1	**0.012**
Afipia | CD	37	2.7E-07	1.07E-06	0.913	**0.012**	1

Appendix XIII: Right colon core OTUs (0.75) that vary significantly in abundance in at least one of the three pairwise comparisons performed (controls vs. CD; controls vs. UC; and CD vs. UC). Kruskal-Wallis followed by Bonferroni correction for multiple comparisons. Values in bold indicate $p <$ Bonferroni corrected significance level of 0.0167. Observations indicate the number of samples from each group employed in the analysis. Mean and std. deviation indicate the mean and std. deviation of the 16S reads obtained. *p*| Control, *p*| CD and *p*| UC indicate the *p* values obtained by comparison to the control, CD and UC, respectively.

OTU | Variable	Taxonomy	Observations	Mean	Std. deviation	*p* | Control	*p* | UC	*p* | CD
	Family						
182122 | CD	Lachnospiraceae	37	0.001382	0.002952	0.077	0.427	1
182122 | Control	Lachnospiraceae	21	0.003299	0.006927	1	**0.034**	0.077
182122 | UC	Lachnospiraceae	14	0.001037	0.002835	**0.034**	1	0.427
183879 | CD	Lachnospiraceae	37	0.000441	0.000974	**0.005**	0.533	1
183879 | Control	Lachnospiraceae	21	0.004161	0.01124	1	0.097	**0.005**
183879 | UC	Lachnospiraceae	14	0.004752	0.013976	0.097	1	0.533
196731 | CD	Lachnospiraceae	37	0.001334	0.003867	0.346	**< 0.0001**	1
196731 | Control	Lachnospiraceae	21	0.000349	0.000478	1	**< 0.0001**	0.346
196731 | UC	Lachnospiraceae	14	0	0	**< 0.0001**	1	**< 0.0001**
303772 | CD	Lachnospiraceae	37	0.000388	0.001227	0.177	**0.007**	1
303772 | Control	Lachnospiraceae	21	0.017307	0.065093	1	0.163	0.177
303772 | UC	Lachnospiraceae	14	0.002921	0.008973	0.163	1	**0.007**
	Genus						
138006 | CD	Alistipes	37	0.00159	0.003173	0.229	**0.006**	1
138006 | Control	Alistipes	21	0.000929	0.001801	1	0.117	0.229
138006 | UC	Alistipes	14	0.000387	0.000823	0.117	1	**0.006**
171559 | CD	Bacteroides	37	0	0	**< 0.0001**	1.000	1
171559 | Control	Bacteroides	21	0.00022	0.000403	1	**< 0.0001**	**< 0.0001**
171559 | UC	Bacteroides	14	0	0	**< 0.0001**	1	1.000
362168 | CD	Bacteroides	37	0.0039	0.01205	**0.039**	0.261	1
362168 | Control	Bacteroides	21	0.001419	0.00439	1	0.540	**0.039**
362168 | UC	Bacteroides	14	0.003015	0.00921	0.540	1	0.261
177005 | CD	Blautia	37	0.00174	0.008173	**0.022**	0.388	1
177005 | Control	Blautia	21	0.003182	0.006608	1	0.304	**0.022**
177005 | UC	Blautia	14	0.001658	0.004534	0.304	1	0.388
514611 | CD	Clostridium	37	0.012226	0.05859	**0.019**	0.633	1
514611 | Control	Clostridium	21	0.005613	0.019147	1	0.154	**0.019**
514611 | UC	Clostridium	14	0.001199	0.002308	0.154	1	0.633
566952 | CD	Clostridium	37	0.000212	0.00033	0.137	0.137	1

566952 | Control	Clostridium	21	0.000305	0.000343	1	**0.011**	0.137
566952 | UC	Clostridium	14	8.29E-05	0.000203	**0.011**	1	0.137
188900 | CD	Faecalibacterium	37	0.016599	0.040401	**0.042**	0.585	1
188900 | Control	Faecalibacterium	21	0.002552	0.00569	1	0.266	**0.042**
188900 | UC	Faecalibacterium	14	0.015322	0.0461	0.266	1	0.585
288565 | CD	Fusobacterium	37	0	0	**< 0.0001**	**< 0.0001**	1
288565 | Control	Fusobacterium	21	0.037348	0.115547	1	0.343	**< 0.0001**
288565 | UC	Fusobacterium	14	0.005165	0.013355	0.343	1	**< 0.0001**
64396 | CD	Fusobacterium	37	0.012184	0.054449	**0.013**	0.993	1
64396 | Control	Fusobacterium	21	0.000174	0.000371	1	**0.050**	**0.013**
64396 | UC	Fusobacterium	14	0.007785	0.022098	**0.050**	1	0.993
541301 | CD	Parabacteroides	37	0.002333	0.00488	0.332	**0.038**	1
541301 | Control	Parabacteroides	21	0.005847	0.015713	1	0.265	0.332
541301 | UC	Parabacteroides	14	0.000801	0.002444	0.265	1	**0.038**
145149 | CD	Veillonella	37	0.001624	0.003682	0.170	0.119	1
145149 | Control	Veillonella	21	0.000151	0.000215	1	**0.012**	0.170
145149 | UC	Veillonella	14	0.008146	0.012085	**0.012**	1	0.119
585419 | CD	Veillonella	37	0.002588	0.006947	0.114	0.059	1
585419 | Control	Veillonella	21	0.000222	0.000403	1	**0.003**	0.114
585419 | UC	Veillonella	14	0.009801	0.018413	**0.003**	1	0.059
	Species						
469709 | CD	Bacteroides dorei	37	0.045166	0.085588	**0.035**	0.529	1
469709 | Control	Bacteroides dorei	21	0.025122	0.065879	1	0.274	**0.035**
469709 | UC	Bacteroides dorei	14	0.028193	0.068693	0.274	1	0.529
258691 | CD	Bacteroides ovatus	37	0.006297	0.015308	**0.038**	0.122	1
258691 | Control	Bacteroides ovatus	21	0.00455	0.017694	1	0.812	**0.038**
258691 | UC	Bacteroides ovatus	14	0.012651	0.046075	0.812	1	0.122
261912 | CD	Dorea formicigenerans	37	0.020078	0.035547	0.540	0.091	1
261912 | Control	Dorea formicigenerans	21	0.015364	0.016683	1	**0.043**	0.540
261912 | UC	Dorea formicigenerans	14	0.003692	0.005013	**0.043**	1	0.091
470973 | CD	Ruminococcus torques	37	0.017474	0.041898	0.637	**0.029**	1
470973 | Control	Ruminococcus torques	21	0.010924	0.03274	1	0.106	0.637
470973 | UC	Ruminococcus torques	14	0.009183	0.029358	0.106	1	**0.029**

Appendix XIV: Right colon taxa that vary significantly in abundance in CD patients in at least one of the three pairwise comparisons performed (mild vs. moderate; mild vs. severe; and moderate vs. severe). Kruskal-Wallis followed by Bonferroni correction for multiple comparisons. Values in bold indicate P < (Bonferroni corrected significance level of 0.0167). Observations indicate the number of samples from each group employed in the analysis. Mean and std. deviation indicate the mean and std. deviation of 16S reads obtained. *p*| mild, *p*| moderate and *p*| severe indicate the *p* values obtained by comparison to the mild, moderate and and severe CD, respectively.

Variable	Observations	Mean	Std. deviation	*p* | mild	*p* | severe	*p* | moderate
PHYLUM						
None						
CLASS						
Clostridia | Mild	9	0.625784	0.174924	1	**0.012**	0.105
Clostridia | Severe	23	0.396447	0.211817	**0.012**	1	0.860
Clostridia | Moderate	5	0.413028	0.180599	0.105	0.860	1
Betaproteobacteria | Mild	9	0.009281	0.018981	1	0.170	**0.013**
Betaproteobacteria | Severe	23	0.021757	0.030244	0.170	1	0.084
Betaproteobacteria | Moderate	5	0.096438	0.1113	**0.013**	0.084	1
ORDER						
Clostridiales | Mild	9	0.625784	0.174924	1	**0.012**	0.105
Clostridiales | Severe	23	0.39644	0.211818	**0.012**	1	0.860
Clostridiales | Moderate	5	0.413028	0.180599	0.105	0.860	1
FAMILY						
Staphylococcaceae | Mild	9	0.00051	0.001014	1	0.025	**0.002**
Staphylococcaceae | Severe	23	0.000515	0.0023	0.025	1	0.089
Staphylococcaceae | Moderate	5	1.6E-06	1.67E-06	**0.002**	0.089	1
Propionibacteriaceae | Mild	9	1.11E-06	1.05E-06	1	**0.016**	0.484
Propionibacteriaceae | Severe	23	6.73E-05	0.000177	**0.016**	1	**0.007**
Propionibacteriaceae | Moderate	5	4E-07	8.94E-07	0.484	**0.007**	1
Acidaminococcaceae | Mild	9	2.22E-06	3.38E-06	1	**0.003**	0.030
Acidaminococcaceae | Severe	23	0.004953	0.015445	**0.003**	1	0.965
Acidaminococcaceae | Moderate	5	0.003424	0.007473	0.030	0.965	1
Bacillaceae | Mild	9	0.000032	9E-05	1	0.137	0.281
Bacillaceae | Severe	23	0.000854	0.002679	0.137	1	**0.016**
Bacillaceae | Moderate	5	1.2E-06	2.68E-06	0.281	**0.016**	1
Carnobacteriaceae | Mild	9	0.00031	0.000459	1	0.206	0.148
Carnobacteriaceae | Severe	23	0.007448	0.031947	0.206	1	**0.008**
Carnobacteriaceae | Moderate	5	5.24E-05	6.03E-05	0.148	**0.008**	1

Sutterellaceae | Mild	9	0.000274	0.000765	1	**0.004**	**0.004**
Sutterellaceae | Severe	23	0.014011	0.030365	**0.004**	1	0.342
Sutterellaceae | Moderate	5	0.08561	0.118934	**0.004**	0.342	1
GENUS						
Atopobium | Mild	9	0.000122	0.000162	1	0.704	0.056
Atopobium | Severe	23	0.000237	0.000527	0.704	1	**0.014**
Atopobium | Moderate	5	9.6E-06	5.55E-06	0.056	**0.014**	1
Propionibacterium | Mild	9	1.11E-06	1.05E-06	1	**0.016**	0.484
Propionibacterium | Severe	23	6.73E-05	0.000177	**0.016**	1	**0.007**
Propionibacterium | Moderate	5	4E-07	8.94E-07	0.484	**0.007**	1
Trichococcus | Mild	9	1.11E-06	2.03E-06	1	**0.002**	0.031
Trichococcus | Severe	23	0	0	**0.002**	1	1.000
Trichococcus | Moderate	5	0	0	0.031	1.000	1
Pectobacterium | Mild	9	0	0	1	1.000	0.029
Pectobacterium | Severe	23	0	0	1.000	1	**0.014**
Pectobacterium | Moderate	5	0.000032	7.16E-05	0.029	**0.014**	1
Granulicatella | Mild	9	0.000291	0.00043	1	0.229	0.167
Granulicatella | Severe	23	0.006988	0.030869	0.229	1	**0.012**
Granulicatella | Moderate	5	0.00005	6.08E-05	0.167	**0.012**	1
Jonquetella | Mild	9	0	0	1	1.000	0.029
Jonquetella | Severe	23	0	0	1.000	1	**0.014**
Jonquetella | Moderate	5	3.32E-05	7.42E-05	0.029	**0.014**	1
Riemerella | Mild	9	0	0	1	1.000	**0.002**
Riemerella | Severe	23	0	0	1.000	1	**0.000**
Riemerella | Moderate	5	1.6E-06	2.61E-06	**0.002**	**0.000**	1
Mogibacterium | Mild	9	1.51E-05	2.99E-05	1	0.187	0.207
Mogibacterium | Severe	23	6.9E-05	0.000164	0.187	1	**0.013**
Mogibacterium | Moderate	5	4E-07	8.94E-07	0.207	**0.013**	1
Staphylococcus | Mild	9	0.000453	0.000856	1	0.024	**0.002**
Staphylococcus | Severe	23	0.000513	0.0023	0.024	1	0.090
Staphylococcus | Moderate	5	1.6E-06	1.67E-06	**0.002**	0.090	1
Sutterella | Mild	9	0.000274	0.000765	1	**0.004**	**0.004**
Sutterella | Severe	23	0.014011	0.030365	**0.004**	1	0.342
Sutterella | Moderate	5	0.08561	0.118934	**0.004**	0.342	1
Phascolarctobacterium | Mild	9	2.22E-06	3.38E-06	1	**0.003**	0.030
Phascolarctobacterium | Severe	23	0.004953	0.015445	**0.003**	1	0.965
Phascolarctobacterium | Moderate	5	0.003424	0.007473	0.030	0.965	1
Comamonas | Mild	9	1.33E-06	2.24E-06	1	**0.016**	0.054
Comamonas | Severe	23	8.7E-08	4.17E-07	**0.016**	1	0.792
Comamonas | Moderate	5	0	0	0.054	0.792	1
Hylemonella | Mild	9	0	0	1	1.000	0.029
Hylemonella | Severe	23	0	0	1.000	1	**0.014**
Hylemonella | Moderate	5	4E-07	8.94E-07	0.029	**0.014**	1
Xenorhabdus | Mild	9	0	0	1	1.000	0.029

Xenorhabdus | Severe	23	0	0	1.000	1	**0.014**
Xenorhabdus | Moderate	5	4E-07	8.94E-07	0.029	**0.014**	1
Averyella | Mild	9	0	0	1	1.000	0.029
Averyella | Severe	23	0	0	1.000	1	**0.014**
Averyella | Moderate	5	4.4E-06	9.84E-06	0.029	**0.014**	1

Appendix XV: Right colon taxa that vary significantly in abundance in UC patients in at least one of the three pairwise comparisons performed (mild vs. moderate; mild vs. severe; and moderate vs. severe). Kruskal-Wallis followed by Bonferroni correction for multiple comparisons. Values in bold indicate $p <$ Bonferroni corrected significance level of 0.0167. Observations indicate the number of samples from each group employed in the analysis. Mean and std. deviation indicate the mean and std. deviation of 16S reads obtained. *p*| mild, *p*| moderate and *p*| severe indicate the *p* values obtained by comparison to the mild, moderate and and severe UC, respectively.

Variable	Observations	Mean	Std. deviation	*p* | mild	*p* | moderate	*p* | severe
Class						
Flavobacteria | mild	3	0.0000000	0.0000000	1	0.588	**0.005**
Flavobacteria | moderate	6	0.0000003	0.0000005	0.588	1	**0.006**
Flavobacteria | Severe	5	0.0000352	0.0000318	**0.005**	**0.006**	1
Order						
Flavobacteriales | mild	3	0.0000000	0.0000000	1	0.588	**0.005**
Flavobacteriales | moderate	6	0.0000003	0.0000005	0.588	1	**0.006**
Flavobacteriales | Severe	5	0.0000352	0.0000318	**0.005**	**0.006**	1
Family						
ClostridialesFamilyXI.IncertaeSedis | mild	3	0.0000008	0.0000014	1	0.032	**0.011**
ClostridialesFamilyXI.IncertaeSedis | moderate	6	0.0000938	0.0001631	0.032	1	0.562
ClostridialesFamilyXI.IncertaeSedis | Severe	5	0.0023648	0.0051211	**0.011**	0.562	1
Flavobacteriaceae | mild	3	0.0000000	0.0000000	1	0.588	**0.005**
Flavobacteriaceae | moderate	6	0.0000003	0.0000005	0.588	1	**0.006**
Flavobacteriaceae | Severe	5	0.0000352	0.0000318	**0.005**	**0.006**	1
Peptococcaceae | mild	3	0.0000000	0.0000000	1	**0.014**	0.074
Peptococcaceae | moderate	6	0.0000075	0.0000063	**0.014**	1	0.481
Peptococcaceae | Severe	5	0.0000128	0.0000214	0.074	0.481	1
Genus						
Catonella | mild	3	0.0000002	0.0000003	1	0.827	0.076
Catonella | moderate	6	0.0000005	0.0000013	0.827	1	**0.017**
Catonella | Severe	5	0.0000109	0.0000112	0.076	**0.017**	1
Fusobacterium | mild	3	0.0000346	0.0000321	1	0.028	**0.013**
Fusobacterium | moderate	6	0.0228940	0.0364835	0.028	1	0.664
Fusobacterium | Severe	5	0.0089288	0.0131355	**0.013**	0.664	1
Streptobacillus | mild	3	0.0000008	0.0000010	1	**0.010**	**0.012**
Streptobacillus | moderate	6	0.0000000	0.0000000	**0.010**	1	1.000
Streptobacillus | Severe	5	0.0000000	0.0000000	**0.012**	1.000	1
Species						
Coprococcus catus | mild	3	0.0000011	0.0000010	1	0.107	0.435
Coprococcus catus | moderate	6	0.0006166	0.0009263	0.107	1	**0.005**

Coprococcus catus | Severe	5	0.0000004	0.0000008	0.435	**0.005**	1
Enterobacter hormaechei | mild	3	0.0000000	0.0000000	1	**0.014**	0.074
Enterobacter hormaechei | moderate	6	0.0035917	0.0086473	**0.014**	1	0.481
Enterobacter hormaechei | Severe	5	0.0000218	0.0000340	0.074	0.481	1

Appendix XVI: Terminal Ileum taxa that vary significantly in abundance in at least one of the three pairwise comparisons performed (controls vs. CD; controls vs. UC; and CD vs. UC). Kruskal-Wallis followed by Bonferroni correction for multiple comparisons. Values in bold indicate $p <$ Bonferroni corrected significance level of 0.0167. Observations indicate the number of samples from each group employed in the analysis. Mean and std. deviation indicate the mean and std. deviation of the 16S reads obtained. *p*| CD, *p*| Control and *p*| UC indicate the *p* values obtained by comparison to the CD, control and UC, respectively.

Taxon | Variable	Observations	Mean	Std. deviation	*p* | CD	*p* | Control	*p* | UC
Class						
Flavobacteria | CD	27	0.118	0.211	1	0.436	**0.012**
Flavobacteria | Control	15	0.062	0.142	0.436	1	**0.004**
Flavobacteria | UC	11	0.343	0.327	**0.012**	**0.004**	1
Gammaproteobacteria | CD	27	0.359	0.179	1	**0.016**	0.211
Gammaproteobacteria | Control	15	0.559	0.268	**0.016**	1	0.411
Gammaproteobacteria | UC	11	0.475	0.274	0.211	0.411	1
Mollicutes | CD	27	0.395	0.336	1	**0.009**	0.075
Mollicutes | Control	15	0.144	0.256	**0.009**	1	0.601
Mollicutes | UC	11	0.248	0.343	0.075	0.601	1
Synechococcophycideae | CD	27	0.000	0.000	1	1.000	**0.008**
Synechococcophycideae | Control	15	0.000	0.000	1.000	1	0.017
Synechococcophycideae | UC	11	0.022	0.049	**0.008**	0.017	1
Order						
Deinococcales | CD	27	0.000	0.000	1	1.000	**0.008**
Deinococcales | Control	15	0.000	0.000	1.000	1	0.017
Deinococcales | UC	11	0.044	0.113	**0.008**	0.017	1
Flavobacteriales | CD	27	0.118	0.211	1	0.436	**0.012**
Flavobacteriales | Control	15	0.061	0.142	0.436	1	**0.004**
Flavobacteriales | UC	11	0.343	0.327	**0.012**	**0.004**	1
Pseudanabaenales | CD	27	0.000	0.000	1	1.000	**0.008**
Pseudanabaenales | Control	15	0.000	0.000	1.000	1	0.017
Pseudanabaenales | UC	11	0.022	0.049	**0.008**	0.017	1
RF39 | CD	27	0.393	0.337	1	**0.009**	**0.015**
RF39 | Control	15	0.144	0.256	**0.009**	1	0.950
RF39 | UC	11	0.172	0.289	**0.015**	0.950	1
Family						
Deinococcaceae | CD	27	0.000	0.000	1	1.000	**0.008**
Deinococcaceae | Control	15	0.000	0.000	1.000	1	0.017
Deinococcaceae | UC	11	0.044	0.113	**0.008**	0.017	1
Flavobacteriaceae | CD	27	0.118	0.211	1	0.436	**0.012**
Flavobacteriaceae | Control	15	0.061	0.142	0.436	1	**0.004**

Flavobacteriaceae | UC	11	0.343	0.327	**0.012**	**0.004**	1
Micrococcaceae | CD	27	0.381	0.297	1	0.097	0.052
Micrococcaceae | Control	15	0.219	0.223	0.097	1	**0.002**
Micrococcaceae | UC	11	0.575	0.279	0.052	**0.002**	1
Oleiphilaceae | CD	27	0.335	0.278	1	**0.014**	0.029
Oleiphilaceae | Control	15	0.143	0.271	**0.014**	1	0.988
Oleiphilaceae | UC	11	0.125	0.220	0.029	0.988	1
Paenibacillaceae | CD	27	0.301	0.183	1	**0.005**	**0.009**
Paenibacillaceae | Control	15	0.505	0.244	**0.005**	1	0.932
Paenibacillaceae | UC	11	0.513	0.264	**0.009**	0.932	1
Peptococcaceae | CD	27	0.367	0.270	1	**0.006**	0.890
Peptococcaceae | Control	15	0.139	0.135	**0.006**	1	0.034
Peptococcaceae | UC	11	0.386	0.337	0.890	0.034	1
Pseudanabaenaceae | CD	27	0.000	0.000	1	1.000	**0.008**
Pseudanabaenaceae | Control	15	0.000	0.000	1.000	1	0.017
Pseudanabaenaceae | UC	11	0.022	0.049	**0.008**	0.017	1
Williamsiaceae | CD	27	0.000	0.000	1	1.000	**0.008**
Williamsiaceae | Control	15	0.000	0.000	1.000	1	0.017
Williamsiaceae | UC	11	0.044	0.112	**0.008**	0.017	1
Genus						
Capnocytophaga | CD	27	0.009	0.034	1	0.292	**< 0.0001**
Capnocytophaga | Control	15	0.046	0.105	0.292	1	**0.008**
Capnocytophaga | UC	11	0.243	0.316	**< 0.0001**	**0.008**	1
Carnobacterium | CD	27	0.258	0.182	1	**0.009**	**0.003**
Carnobacterium | Control	15	0.491	0.286	**0.009**	1	0.548
Carnobacterium | UC	11	0.594	0.290	**0.003**	0.548	1
Catonella | CD	27	0.081	0.226	1	0.551	**0.003**
Catonella | Control	15	0.136	0.296	0.551	1	0.026
Catonella | UC	11	0.318	0.313	**0.003**	0.026	1
Deinococcus | CD	27	0.000	0.000	1	1.000	**0.008**
Deinococcus | Control	15	0.000	0.000	1.000	1	0.017
Deinococcus | UC	11	0.044	0.113	**0.008**	0.017	1
Halomicronema | CD	27	0.000	0.000	1	1.000	**0.008**
Halomicronema | Control	15	0.000	0.000	1.000	1	0.017
Halomicronema | UC	11	0.022	0.049	**0.008**	0.017	1
Marinobacterium | CD	27	0.012	0.046	1	**0.000**	0.077
Marinobacterium | Control	15	0.373	0.427	**0.000**	1	0.211
Marinobacterium | UC	11	0.181	0.310	0.077	0.211	1
Riemerella | CD	27	0.000	0.000	1	0.146	**0.012**
Riemerella | Control	15	0.072	0.221	0.146	1	0.277
Riemerella | UC	11	0.101	0.246	**0.012**	0.277	1
Roseburia | CD	27	0.302	0.183	1	0.467	**0.006**
Roseburia | Control	15	0.409	0.289	0.467	1	0.062
Roseburia | UC	11	0.578	0.299	**0.006**	0.062	1
Rothia | CD	27	0.381	0.297	1	0.097	0.061

Rothia | Control	15	0.219	0.221	0.097	1	**0.002**
Rothia | UC	11	0.564	0.288	0.061	**0.002**	1
Species						
Capnocytophagaochracea | CD	27	0.000	0.000	1	0.057	**0.004**
Capnocytophagaochracea | Control	15	0.109	0.248	0.057	1	0.304
Capnocytophagaochracea | UC	11	0.109	0.184	**0.004**	0.304	1
Clostridiumbifermentans | CD	27	0.000	0.000	1	1.000	**0.008**
Clostridiumbifermentans | Control	15	0.000	0.000	1.000	1	0.017
Clostridiumbifermentans | UC	11	0.105	0.277	**0.008**	0.017	1
Clostridiumpiliforme | CD	27	0.000	0.000	1	1.000	**0.008**
Clostridiumpiliforme | Control	15	0.000	0.000	1.000	1	0.017
Clostridiumpiliforme | UC	11	0.092	0.250	**0.008**	0.017	1
Corynebacteriumaurimucosum | CD	27	0.000	0.000	1	1.000	**0.008**
Corynebacteriumaurimucosum | Control	15	0.000	0.000	1.000	1	0.017
Corynebacteriumaurimucosum | UC	11	0.124	0.295	**0.008**	0.017	1
Eubacteriumdolichum | CD	27	0.102	0.247	1	0.147	0.061
Eubacteriumdolichum | Control	15	0.000	0.000	0.147	1	**0.004**
Eubacteriumdolichum | UC	11	0.222	0.285	0.061	**0.004**	1
Granulicatellaadiacens | CD	27	0.271	0.298	1	**0.013**	0.644
Granulicatellaadiacens | Control	15	0.076	0.170	**0.013**	1	0.111
Granulicatellaadiacens | UC	11	0.234	0.318	0.644	0.111	1
Prevotellaveroralis | CD	27	0.328	0.281	1	**0.014**	0.139
Prevotellaveroralis | Control	15	0.136	0.221	**0.014**	1	0.508
Prevotellaveroralis | UC	11	0.181	0.205	0.139	0.508	1
Roseburiafaecis | CD	27	0.057	0.190	1	0.759	**0.014**
Roseburiafaecis | Control	15	0.091	0.229	0.759	1	0.050
Roseburiafaecis | UC	11	0.278	0.336	**0.014**	0.050	1
Streptococcusthermophilus | CD	27	0.465	0.269	1	0.021	0.235
Streptococcusthermophilus | Control	15	0.289	0.182	0.021	1	**0.003**
Streptococcusthermophilus | UC	11	0.540	0.250	0.235	**0.003**	1
Veillonellaparvula | CD	27	0.000	0.000	1	1.000	**0.008**
Veillonellaparvula | Control	15	0.000	0.000	1.000	1	0.017
Veillonellaparvula | UC	11	0.034	0.080	**0.008**	0.017	1
unclassified_Enterobacteriaceae | CD	27	0.332	0.217	1	**0.012**	0.085
unclassified_Enterobacteriaceae | Control	15	0.466	0.183	**0.012**	1	0.625
unclassified_Enterobacteriaceae | UC	11	0.469	0.261	0.085	0.625	1
unclassified_Pasteurellaceae | CD	27	0.318	0.221	1	0.408	**0.008**
unclassified_Pasteurellaceae | Control	15	0.282	0.288	0.408	1	**0.002**
unclassified_Pasteurellaceae | UC	11	0.669	0.301	**0.008**	**0.002**	1
unclassified_Clostridium | CD	27	0.440	0.220	1	0.615	0.020
unclassified_Clostridium | Control	15	0.393	0.268	0.615	1	**0.012**
unclassified_Clostridium | UC	11	0.361	0.139	0.020	**0.012**	1
unclassified_Roseomonas | CD	27	0.000	0.000	1	1.000	**0.008**
unclassified_Roseomonas | Control	15	0.000	0.000	1.000	1	0.017

unclassified_Roseomonas | UC	11	0.044	0.099	**0.008**	0.017	1
unclassified_Streptococcus | CD	27	0.449	0.274	1	0.542	**0.016**
unclassified_Streptococcus | Control	15	0.382	0.199	0.542	1	0.092
unclassified_Streptococcus | UC	11	0.258	0.203	**0.016**	0.092	1

Appendix XVII: Terminal ileum core OTUs (0.75) that vary significantly in abundance in at least one of the three pairwise comparisons performed (controls vs. CD; controls vs. UC; and CD vs. UC). Kruskal-Wallis followed by Bonferroni correction for multiple comparisons. Values in bold indicate $p <$ (Bonferroni corrected significance level of 0.0167). Mean and std. deviation indicate the mean and std. deviation of the 16S reads obtained. *p*| CD, *p*| Control and *p*| UC indicate the *p* values obtained by comparison to the CD, control and UC, respectively.

Variable	Taxonomy	Mean	Std. deviation	*p*|CD	*p*| Control	*p*|UC
165397 | CD	Lachnospiraceae g__; s__	0.000051	0.000120	1	0.055	0.237
165397 | Control	Lachnospiraceae g__; s__	0.000146	0.000371	0.055	1	**0.009**
165397 | UC	Lachnospiraceae g__; s__	0.000025	0.000081	0.237	**0.009**	1
206865 | CD	Lachnospiraceae g__; s__	0.000038	0.000084	1	0.148	0.171
206865 | Control	Lachnospiraceae g__; s__	0.000906	0.002950	0.148	1	**0.016**
206865 | UC	Lachnospiraceae g__; s__	0.000106	0.000327	0.171	**0.016**	1
293896 | CD	Lachnospiraceae g__; s__	0.000155	0.000468	1	0.194	0.089
293896 | Control	Lachnospiraceae g__; s__	0.000032	0.000055	0.194	1	**0.010**
293896 | UC	Lachnospiraceae g__; s__	0.000003	0.000006	0.089	**0.010**	1
296516 | CD	Lachnospiraceae g__; s__	0.000004	0.000009	1	**0.005**	0.490
296516 | Control	Lachnospiraceae g__; s__	0.000047	0.000107	**0.005**	1	0.094
296516 | UC	Lachnospiraceae g__; s__	0.000004	0.000007	0.490	0.094	1
319199 | CD	Lachnospiraceae g__; s__	0.000009	0.000027	1	**0.008**	0.386
319199 | Control	Lachnospiraceae g__; s__	0.000234	0.000361	**0.008**	1	0.162
319199 | UC	Lachnospiraceae g__; s__	0.000098	0.000244	0.386	0.162	1
352304 | CD	Lachnospiraceae g__; s__	0.000266	0.000684	1	**0.011**	0.031
352304 | Control	Lachnospiraceae g__; s__	0.005595	0.011447	**0.011**	1	0.859
352304 | UC	Lachnospiraceae g__; s__	0.001019	0.002115	0.031	0.859	1
453710 | CD	Ruminococcaceae g__; s__	0.000048	0.000110	1	**0.011**	0.248
453710 | Control	Ruminococcaceae g__; s__	0.000005	0.000012	**0.011**	1	0.295
453710 | UC	Ruminococcaceae g__; s__	0.000010	0.000018	0.248	0.295	1
182190 | CD	Clostridium; s__	0.000036	0.000084	1	**0.005**	0.383
182190 | Control	Clostridium; s__	0.000234	0.000418	**0.005**	1	0.122
182190 | UC	Clostridium; s__	0.000291	0.000558	0.383	0.122	1
204180 | CD	Clostridium; s__	0.002768	0.008083	1	0.020	0.481
204180 | Control	Clostridium; s__	0.014063	0.032596	0.020	1	**0.012**
204180 | UC	Clostridium; s__	0.001463	0.003091	0.481	**0.012**	1
237324 | CD	Clostridium; s__	0.001487	0.002935	1	0.632	**0.014**
237324 | Control	Clostridium; s__	0.008490	0.028728	0.632	1	0.073
237324 | UC	Clostridium; s__	0.000482	0.001480	**0.014**	0.073	1
173934 | CD	Coprococcus; s__	0.000078	0.000236	1	**0.015**	0.229
173934 | Control	Coprococcus; s__	0.000167	0.000246	**0.015**	1	0.358
173934 | UC	Coprococcus; s__	0.000100	0.000160	0.229	0.358	1

178146 | CD	Coprococcus; s__	0.000001	0.000003	1	**0.000**	0.020
178146 | Control	Coprococcus; s__	0.000008	0.000019	**0.000**	1	0.427
178146 | UC	Coprococcus; s__	0.000016	0.000038	0.020	0.427	1
288565 | CD	Fusobacterium; s__	0.000029	0.000059	1	**0.015**	0.800
288565 | Control	Fusobacterium; s__	0.016105	0.039962	**0.015**	1	0.078
288565 | UC	Fusobacterium; s__	0.009521	0.023194	0.800	0.078	1
229097 | CD	Ruminococcus; s__	0.000002	0.000004	1	**0.004**	0.902
229097 | Control	Ruminococcus; s__	0.000019	0.000041	**0.004**	1	0.024
229097 | UC	Ruminococcus; s__	0.000013	0.000024	0.902	0.024	1
349002 | CD	Ruminococcus; s__	0.000049	0.000087	1	**0.008**	0.657
349002 | Control	Ruminococcus; s__	0.000138	0.000164	**0.008**	1	0.078
349002 | UC	Ruminococcus; s__	0.000045	0.000056	0.657	0.078	1
145149 | CD	Veillonella; s__	0.003532	0.010049	1	0.045	0.065
145149 | Control	Veillonella; s__	0.000669	0.001722	0.045	1	**0.001**
145149 | UC	Veillonella; s__	0.007863	0.011084	0.065	**0.001**	1
585419 | CD	Veillonella; s__	0.003473	0.008682	1	0.174	0.042
585419 | Control	Veillonella; s__	0.000820	0.001390	0.174	1	**0.004**
585419 | UC	Veillonella; s__	0.009678	0.015268	0.042	**0.004**	1
544854 | CD	Vitreoscilla; s__	0.000004	0.000008	1	0.864	**0.016**
544854 | Control	Vitreoscilla; s__	0.000295	0.000814	0.864	1	0.047
544854 | UC	Vitreoscilla; s__	0.000307	0.000684	**0.016**	0.047	1
322235 | CD	Bacteroidesuniformis	0.007925	0.013561	1	**0.013**	0.706
322235 | Control	Bacteroidesuniformis	0.001072	0.001374	**0.013**	1	0.018
322235 | UC	Bacteroidesuniformis	0.009546	0.018505	0.706	0.018	1
349216 | CD	Bacteroidesuniformis	0.000012	0.000027	1	**0.002**	0.453
349216 | Control	Bacteroidesuniformis	0.000001	0.000002	**0.002**	1	0.068
349216 | UC	Bacteroidesuniformis	0.000006	0.000010	0.453	0.068	1
46789 | CD	Clostridiumperfringens	0.000008	0.000017	1	**0.017**	0.597
46789 | Control	Clostridiumperfringens	0.006435	0.016010	**0.017**	1	0.136
46789 | UC	Clostridiumperfringens	0.003594	0.009424	0.597	0.136	1
261912 | CD	Doreaformicigenerans	0.011727	0.021815	1	0.043	0.387
261912 | Control	Doreaformicigenerans	0.017438	0.019220	0.043	1	**0.016**
261912 | UC	Doreaformicigenerans	0.003587	0.005951	0.387	**0.016**	1
109018 | CD	Eubacteriumrectale	0.000388	0.001252	1	**0.005**	0.387
109018 | Control	Eubacteriumrectale	0.002187	0.005205	**0.005**	1	0.125
109018 | UC	Eubacteriumrectale	0.001722	0.003453	0.387	0.125	1
175682 | CD	Eubacteriumrectale	0.000165	0.000495	1	**0.002**	0.201
175682 | Control	Eubacteriumrectale	0.002891	0.006867	**0.002**	1	0.165
175682 | UC	Eubacteriumrectale	0.003820	0.007857	0.201	0.165	1
353985 | CD	Parabacteroidesdistasonis	0.015896	0.039710	1	**0.006**	**0.015**
353985 | Control	Parabacteroidesdistasonis	0.001977	0.004657	**0.006**	1	0.931
353985 | UC	Parabacteroidesdistasonis	0.003250	0.009898	**0.015**	0.931	1

Appendix XVIII: Terminal ileum taxa that vary significantly in abundance in CD patients in at least one of the three pairwise comparisons performed (mild vs. moderate; mild vs. severe; and moderate vs. severe). Kruskal-Wallis followed by Bonferroni correction for multiple comparisons. Values in bold indicate $p <$ Bonferroni corrected significance level of 0.0167. Observations indicate the number of samples from each group employed in the analysis. Mean and std. deviation indicate the mean and std. deviation of the 16S reads obtained. *p*| mild, *p*| moderate and *p*| severe indicate the *p* values obtained by comparison to the mild, moderate and severe CD, respectively.

Variable	Observations	Mean	Std. deviation	*p* | mild	*p* | moderate	*p* | severe
Class						
Anaerolineae | mild	5	0.0000019	0.0000034	1	0.026	**0.003**
Anaerolineae | moderate	4	0.0000000	0.0000000	0.026	1	1.000
Anaerolineae | Severe	18	0.0000000	0.0000000	**0.003**	1.000	1
Order						
Acidithiobacillales | mild	5	0.0000013	0.0000018	1	0.026	**0.003**
Acidithiobacillales | moderate	4	0.0000000	0.0000000	0.026	1	1.000
Acidithiobacillales | Severe	18	0.0000000	0.0000000	**0.003**	1.000	1
Aeromonadales | mild	5	0.0000015	0.0000025	1	0.394	**0.011**
Aeromonadales | moderate	4	0.0000002	0.0000003	0.394	1	0.192
Aeromonadales | Severe	18	0.0000000	0.0000000	**0.011**	0.192	1
Anaerolineales | mild	5	0.0000019	0.0000034	1	0.026	**0.003**
Anaerolineales | moderate	4	0.0000000	0.0000000	0.026	1	1.000
Anaerolineales | Severe	18	0.0000000	0.0000000	**0.003**	1.000	1
Desulfuromonadales | mild	5	0.0000004	0.0000005	1	0.026	**0.003**
Desulfuromonadales | moderate	4	0.0000000	0.0000000	0.026	1	1.000
Desulfuromonadales | Severe	18	0.0000000	0.0000000	**0.003**	1.000	1
Rhodobacterales | mild	5	0.0000023	0.0000037	1	**0.013**	**0.003**
Rhodobacterales | moderate	4	0.0000000	0.0000000	**0.013**	1	0.790
Rhodobacterales | Severe	18	0.0000001	0.0000002	**0.003**	0.790	1
Family						
Acidithiobacillaceae | mild	5	0.0000013	0.0000018	1	0.026	**0.003**
Acidithiobacillaceae | moderate	4	0.0000000	0.0000000	0.026	1	1.000
Acidithiobacillaceae | Severe	18	0.0000000	0.0000000	**0.003**	1.000	1
Aeromonadaceae | mild	5	0.0000015	0.0000025	1	0.394	**0.011**
Aeromonadaceae | moderate	4	0.0000002	0.0000003	0.394	1	0.192
Aeromonadaceae | Severe	18	0.0000000	0.0000000	**0.011**	0.192	1
Anaerolinaceae | mild	5	0.0000019	0.0000034	1	0.026	**0.003**
Anaerolinaceae | moderate	4	0.0000000	0.0000000	0.026	1	1.000
Anaerolinaceae | Severe	18	0.0000000	0.0000000	**0.003**	1.000	1

Geobacteraceae | mild	5	0.0000004	0.0000005	1	0.026	**0.003**
Geobacteraceae | moderate	4	0.0000000	0.0000000	0.026	1	1.000
Geobacteraceae | Severe	18	0.0000000	0.0000000	**0.003**	1.000	1
Hyphomicrobiaceae | mild	5	0.0000006	0.0000009	1	0.026	**0.003**
Hyphomicrobiaceae | moderate	4	0.0000000	0.0000000	0.026	1	1.000
Hyphomicrobiaceae | Severe	18	0.0000000	0.0000000	**0.003**	1.000	1
Rhodobacteraceae | mild	5	0.0000023	0.0000037	1	**0.013**	**0.003**
Rhodobacteraceae | moderate	4	0.0000000	0.0000000	**0.013**	1	0.790
Rhodobacteraceae | Severe	18	0.0000001	0.0000002	**0.003**	0.790	1
Genus						
Acidithiobacillus | mild	5	0.0000013	0.0000018	1	0.026	**0.003**
Acidithiobacillus | moderate	4	0.0000000	0.0000000	0.026	1	1.000
Acidithiobacillus | Severe	18	0.0000000	0.0000000	**0.003**	1.000	1
Acidovorax | mild	5	0.0000074	0.0000129	1	0.064	**0.005**
Acidovorax | moderate	4	0.0000005	0.0000010	0.064	1	0.760
Acidovorax | Severe	18	0.0000062	0.0000259	**0.005**	0.760	1
Aeromonas | mild	5	0.0000011	0.0000017	1	0.394	**0.011**
Aeromonas | moderate	4	0.0000002	0.0000003	0.394	1	0.192
Aeromonas | Severe	18	0.0000000	0.0000000	**0.011**	0.192	1
Devosia | mild	5	0.0000006	0.0000009	1	0.026	**0.003**
Devosia | moderate	4	0.0000000	0.0000000	0.026	1	1.000
Devosia | Severe	18	0.0000000	0.0000000	**0.003**	1.000	1
Geobacter | mild	5	0.0000004	0.0000005	1	0.026	**0.003**
Geobacter | moderate	4	0.0000000	0.0000000	0.026	1	1.000
Geobacter | Severe	18	0.0000000	0.0000000	**0.003**	1.000	1
Laribacter | mild	5	0.0000007	0.0000010	1	0.026	**0.003**
Laribacter | moderate	4	0.0000000	0.0000000	0.026	1	1.000
Laribacter | Severe	18	0.0000000	0.0000000	**0.003**	1.000	1
Luteibacter | mild	5	0.0000013	0.0000018	1	0.394	**0.011**
Luteibacter | moderate	4	0.0000002	0.0000003	0.394	1	0.192
Luteibacter | Severe	18	0.0000000	0.0000000	**0.011**	0.192	1
Mobiluncus | mild	5	0.0000143	0.0000316	1	0.486	**0.014**
Mobiluncus | moderate	4	0.0000003	0.0000005	0.486	1	0.159
Mobiluncus | Severe	18	0.0000000	0.0000000	**0.014**	0.159	1
Olsenella | mild	5	0.0000000	0.0000000	1	**0.016**	0.751
Olsenella | moderate	4	0.0000021	0.0000029	**0.016**	1	**0.008**
Olsenella | Severe	18	0.0000001	0.0000003	0.751	**0.008**	1
Rhodobacter | mild	5	0.0000005	0.0000005	1	**0.005**	**0.000**
Rhodobacter | moderate	4	0.0000000	0.0000000	**0.005**	1	1.000
Rhodobacter | Severe	18	0.0000000	0.0000000	**0.000**	1.000	1
T78 | mild	5	0.0000019	0.0000034	1	0.026	**0.003**
T78 | moderate	4	0.0000000	0.0000000	0.026	1	1.000
T78 | Severe	18	0.0000000	0.0000000	**0.003**	1.000	1
Species						

Acidithiobacillusalbertensis | mild	5	0.0000013	0.0000018	1	0.026	**0.003**
Acidithiobacillusalbertensis | moderate	4	0.0000000	0.0000000	0.026	1	1.000
Acidithiobacillusalbertensis | Severe	18	0.0000000	0.0000000	**0.003**	1.000	1
Actinomyceshyovaginalis | mild	5	0.0000335	0.0000734	1	**0.015**	**0.004**
Actinomyceshyovaginalis | moderate	4	0.0000000	0.0000000	**0.015**	1	0.774
Actinomyceshyovaginalis | Severe	18	0.0000014	0.0000058	**0.004**	0.774	1
Bacteroidesuniformis | mild	5	0.0006409	0.0005348	1	**0.007**	0.094
Bacteroidesuniformis | moderate	4	0.0153539	0.0174177	**0.007**	1	0.082
Bacteroidesuniformis | Severe	18	0.0084998	0.0142248	0.094	0.082	1
Blautiaproducta | mild	5	0.0006840	0.0008356	1	0.044	**0.013**
Blautiaproducta | moderate	4	0.0000292	0.0000430	0.044	1	0.859
Blautiaproducta | Severe	18	0.0000366	0.0000529	**0.013**	0.859	1
Clostridiumdifficile | mild	5	0.0000000	0.0000000	1	**0.003**	0.610
Clostridiumdifficile | moderate	4	0.0001805	0.0003463	**0.003**	1	**0.002**
Clostridiumdifficile | Severe	18	0.0000008	0.0000027	0.610	**0.002**	1
Haemophilusparainfluenzae | mild	5	0.0000000	0.0000000	1	**0.004**	0.146
Haemophilusparainfluenzae | moderate	4	0.0000044	0.0000048	**0.004**	1	0.028
Haemophilusparainfluenzae | Severe	18	0.0000014	0.0000025	0.146	0.028	1
Laribacterhongkongensis | mild	5	0.0000007	0.0000010	1	0.026	**0.003**
Laribacterhongkongensis | moderate	4	0.0000000	0.0000000	0.026	1	1.000
Laribacterhongkongensis | Severe	18	0.0000000	0.0000000	**0.003**	1.000	1
Luteibacterrhizovicinus | mild	5	0.0000013	0.0000018	1	0.394	**0.011**
Luteibacterrhizovicinus | moderate	4	0.0000002	0.0000003	0.394	1	0.192
Luteibacterrhizovicinus | Severe	18	0.0000000	0.0000000	**0.011**	0.192	1
Mobiluncuscurtisii | mild	5	0.0000143	0.0000316	1	0.486	**0.014**
Mobiluncuscurtisii | moderate	4	0.0000003	0.0000005	0.486	1	0.159
Mobiluncuscurtisii | Severe	18	0.0000000	0.0000000	**0.014**	0.159	1
Selenomonasruminantium | mild	5	0.0000007	0.0000010	1	0.026	**0.003**
Selenomonasruminantium | moderate	4	0.0000000	0.0000000	0.026	1	1.000
Selenomonasruminantium | Severe	18	0.0000000	0.0000000	**0.003**	1.000	1
Sphingomonasazotifigens | mild	5	0.0000007	0.0000010	1	0.026	**0.003**
Sphingomonasazotifigens | moderate	4	0.0000000	0.0000000	0.026	1	1.000
Sphingomonasazotifigens | Severe	18	0.0000000	0.0000000	**0.003**	1.000	1
unclassified_Bifidobacterium | mild	5	0.0000000	0.0000000	1	**0.011**	0.341
unclassified_Bifidobacterium | moderate	4	0.0000364	0.0000675	**0.011**	1	0.028
unclassified_Bifidobacterium | Severe	18	0.0000440	0.0001842	0.341	0.028	1
unclassified_Corynebacterium | mild	5	0.0000204	0.0000419	1	**0.016**	0.036
unclassified_Corynebacterium | moderate	4	0.0000000	0.0000000	**0.016**	1	0.308
unclassified_Corynebacterium | Severe	18	0.0000211	0.0000851	0.036	0.308	1
unclassified_Fibrobacter | mild	5	0.0000007	0.0000013	1	0.026	**0.003**
unclassified_Fibrobacter | moderate	4	0.0000000	0.0000000	0.026	1	1.000
unclassified_Fibrobacter | Severe	18	0.0000000	0.0000000	**0.003**	1.000	1

Appendix XIX: Terminal ileum taxa that vary significantly in abundance in UC patients in at least one of the three pairwise comparisons performed (mild vs. moderate; mild vs. severe; and moderate vs. severe). Kruskal-Wallis followed by Bonferroni correction for multiple comparisons. Values in bold indicate $p <$ Bonferroni corrected significance level of 0.0167. Observations indicate the number of samples from each group employed in the analysis. Mean and std. deviation indicate the mean and std. deviation of the 16S reads obtained. *p*| mild, *p*| moderate and *p*| severe indicate the *p* values obtained by comparison to the mild, moderate and severe UC, respectively.

Variable	Observations	Mean	Std. deviation	*p* | mild	*p* | moderate	*p* | severe
Genus						
Capnocytophaga | mild	2	0.0000019	0.0000026	1	0.568	0.154
Capnocytophaga | moderate	5	0.0000003	0.0000005	0.568	1	**0.011**
Capnocytophaga | Severe	4	0.0000439	0.0000646	0.154	**0.011**	1
Turicibacter | mild	2	0.0000181	0.0000124	1	0.050	**0.009**
Turicibacter | moderate	5	0.0000009	0.0000016	0.050	1	0.349
Turicibacter | Severe	4	0.0000000	0.0000000	**0.009**	0.349	1
Species						
Bifidobacteriumadolescentis | mild	2	0.0000054	0.0000050	1	**0.006**	0.031
Bifidobacteriumadolescentis | moderate	5	0.0000000	0.0000000	**0.006**	1	0.520
Bifidobacteriumadolescentis | Severe	4	0.0000003	0.0000005	0.031	0.520	1
Capnocytophagaochracea | mild	2	0.0000000	0.0000000	1	1.000	0.027
Capnocytophagaochracea | moderate	5	0.0000000	0.0000000	1.000	1	**0.004**
Capnocytophagaochracea | Severe	4	0.0000106	0.0000155	0.027	**0.004**	1

Appendix XX: Microbial metabolic pathways with significant differential abundance in the LC communities of new onset pediatric IBD patients. Relative abundance of KEGG pathways were predicted by PICRUST approach [244]. Statistical significance was assessed using a Kruskal-Wallis test with a Dunn's post hoc test. Values in bold indicate $p <$ Bonferroni corrected significance level of 0.0167. Observations indicate the number of samples from each group employed in the analysis. Mean and std. deviation indicate the mean and std. deviation of the 16S reads obtained. *p*| CD, *p*| Control and *p*| UC indicate the *p* values obtained by comparison to the CD, control and UC, respectively.

Pathway|Variable	Obser-vations	Mean	Std. deviation	*p*|CD	*p*| Control	*p*|UC
Bacterial secretion system | CD	39	0.005634	0.000757	1	**0.002**	0.338
Bacterial secretion system | Control	34	0.005241	0.000812	**0.002**	1	0.217
Bacterial secretion system | UC	13	0.005503	0.000861	0.338	0.217	1
Carbohydrate metabolism | CD	39	0.001836	0.00035	1	**0.004**	0.263
Carbohydrate metabolism | Control	34	0.00208	0.000346	**0.004**	1	0.324
Carbohydrate metabolism | UC	13	0.001968	0.000382	0.263	0.324	1
Carotenoid biosynthesis | CD	39	1.54E-05	2.59E-05	1	**0.008**	0.822
Carotenoid biosynthesis | Control	34	6.94E-06	1.07E-05	**0.008**	1	0.094
Carotenoid biosynthesis | UC	13	2.99E-05	5.32E-05	0.822	0.094	1
D-Alanine metabolism | CD	39	0.001019	0.000312	1	**0.016**	0.290
D-Alanine metabolism | Control	34	0.0009	8.14E-05	**0.016**	1	0.483
D-Alanine metabolism | UC	13	0.000968	0.000199	0.290	0.483	1
DNA replication | CD	39	0.006332	0.000679	1	0.098	0.218
DNA replication | Control	34	0.00607	0.000554	0.098	1	**0.016**
DNA replication | UC	13	0.006459	0.000573	0.218	**0.016**	1
Fc epsilon RI signaling pathway | CD	39	0.000126	6.83E-05	1	0.133	**0.005**
Fc epsilon RI signaling pathway | Control	34	0.000107	7.68E-05	0.133	1	0.091
Fc epsilon RI signaling pathway | UC	13	7.29E-05	8.37E-05	**0.005**	0.091	1
Glycolysis / Gluconeogenesis | CD	39	0.000346	0.000117	1	**0.010**	0.223
Glycolysis / Gluconeogenesis | Control	34	0.000417	0.000122	**0.010**	1	0.504
Glycolysis / Gluconeogenesis | UC	13	0.000398	0.000127	0.223	0.504	1
Glycosphingolipid biosynthesis - lacto and neolacto series | CD	39	0.009058	0.00089	1	0.113	0.118
Glycosphingolipid biosynthesis - lacto and neolacto series | Control	34	0.008762	0.000752	0.113	1	**0.008**
Glycosphingolipid biosynthesis - lacto and neolacto series | UC	13	0.009257	0.000638	0.118	**0.008**	1
Oocyte meiosis | CD	39	0.011739	0.000916	1	0.571	0.033
Oocyte meiosis | Control	34	0.011624	0.000778	0.571	1	**0.012**

Oocyte meiosis | UC	13	0.012169	0.001287	0.033	**0.012**	1
Pantothenate and CoA biosynthesis | CD	39	0.008776	0.000759	1	**0.011**	0.803
Pantothenate and CoA biosynthesis | Control	34	0.009249	0.00083	**0.011**	1	0.113
Pantothenate and CoA biosynthesis | UC	13	0.008811	0.000674	0.803	0.113	1
Proteasome | CD	39	0.021282	0.001484	1	0.298	0.077
Proteasome | Control	34	0.020833	0.001069	0.298	1	**0.013**
Proteasome | UC	13	0.021738	0.00114	0.077	**0.013**	1
Protein digestion and absorption | CD	39	0.017667	0.001567	1	0.116	0.192
Protein digestion and absorption | Control	34	0.017227	0.001254	0.116	1	**0.016**
Protein digestion and absorption | UC	13	0.017865	0.001141	0.192	**0.016**	1
Protein processing in endoplasmic reticulum | CD	39	0.004618	0.000403	1	0.084	0.191
Protein processing in endoplasmic reticulum | Control	34	0.004464	0.00042	0.084	1	**0.011**
Protein processing in endoplasmic reticulum | UC	13	0.004762	0.000401	0.191	**0.011**	1
Pyruvate metabolism | CD	39	0.008561	0.00171	1	**0.016**	0.021
Pyruvate metabolism | Control	34	0.007708	0.001488	**0.016**	1	0.600
Pyruvate metabolism | UC	13	0.007625	0.001831	0.021	0.600	1
Retinol metabolism | CD	39	0.000455	4.63E-05	1	**0.009**	0.700
Retinol metabolism | Control	34	0.000431	4.91E-05	**0.009**	1	0.134
Retinol metabolism | UC	13	0.000461	5.56E-05	0.700	0.134	1
Viral myocarditis | CD	39	0.010543	0.001043	1	**0.012**	0.977
Viral myocarditis | Control	34	0.011435	0.001585	**0.012**	1	0.065

Appendix XXI: Microbial metabolic pathways with significant differential abundance in the RC communities of new onset pediatric IBD patients. Relative abundance of KEGG pathways were predicted by PICRUST approach [244]. Statistical significance was assessed using a Kruskal-Wallis test with a Dunn's post hoc test. Values in bold indicate $p <$ Bonferroni corrected significance level of 0.0167. Observations indicate the number of samples from each group employed in the analysis. Mean and std. deviation indicate the mean and std. deviation of the 16S reads obtained. *p*| CD, *p*| Control and *p*| UC indicate the *p* values obtained by comparison to the CD, control and UC, respectively.

Pathway|Variable	Observations	Mean	Std. deviation	*p*|CD	*p*| Control	*p*|UC
1-1-1-Trichloro-2-2-bis(4-chlorophenyl)ethane (DDT) degradation | CD	37	1.081E-05	2.27E-05	1	**0.014**	0.063
1-1-1-Trichloro-2-2-bis(4-chlorophenyl)ethane (DDT) degradation | Control	21	2.117E-06	5.16681E-06	**0.014**	1	0.794
1-1-1-Trichloro-2-2-bis(4-chlorophenyl)ethane (DDT) degradation | UC	14	4.96E-06	1.20397E-05	0.063	0.794	1
Aminobenzoate degradation | CD	37	0.0012117	0.000473121	1	0.020	0.273
Aminobenzoate degradation | Control	21	0.0009349	0.000200239	0.020	1	**0.005**
Aminobenzoate degradation | UC	14	0.0012601	0.000319966	0.273	**0.005**	1
Amyotrophic lateral sclerosis (ALS) | CD	37	0.0001602	0.000125513	1	0.261	**0.016**
Amyotrophic lateral sclerosis (ALS) | Control	21	0.000112	5.7639E-05	0.261	1	**0.002**
Amyotrophic lateral sclerosis (ALS) | UC	14	0.0002392	0.000120336	**0.016**	**0.002**	1
Arachidonic acid metabolism | CD	37	0.0001698	0.000128978	1	0.226	0.083
Arachidonic acid metabolism | Control	21	0.000121	9.28389E-05	0.226	1	**0.011**
Arachidonic acid metabolism | UC	14	0.0002385	0.00015256	0.083	**0.011**	1
Bacterial motility proteins | CD	37	0.0039253	0.002665726	1	0.375	**0.011**
Bacterial motility proteins | Control	21	0.0047393	0.003105347	0.375	1	0.109
Bacterial motility proteins | UC	14	0.0063169	0.003527738	**0.011**	0.109	1
Bacterial secretion system | CD	37	0.0056243	0.000814231	1	**0.009**	0.928
Bacterial secretion system | Control	21	0.0050945	0.000331922	**0.009**	1	0.032
Bacterial secretion system | UC	14	0.0059281	0.001451929	0.928	0.032	1
Biosynthesis and biodegradation of secondary metabolites | CD	37	0.0007948	0.000184165	1	0.352	**0.007**
Biosynthesis and biodegradation of secondary metabolites | Control	21	0.0007318	0.000167323	0.352	1	**0.002**
Biosynthesis and biodegradation of secondary metabolites | UC	14	0.0010134	0.000366905	**0.007**	**0.002**	1
Biosynthesis of ansamycins | CD	37	0.0010305	0.000196751	1	**0.007**	0.373

Biosynthesis of ansamycins | Control	21	0.0012147	0.000273622	**0.007**	1	0.185
Biosynthesis of ansamycins | UC	14	0.0011138	0.000347969	0.373	0.185	1
Carbohydrate digestion and absorption | CD	37	0.0001343	0.000162171	1	0.212	0.077
Carbohydrate digestion and absorption | Control	21	7.694E-05	8.2554E-05	0.212	1	**0.009**
Carbohydrate digestion and absorption | UC	14	0.0002022	0.00014089	0.077	**0.009**	1
Cardiac muscle contraction | CD	37	3.753E-06	1.04284E-05	1	0.073	0.137
Cardiac muscle contraction | Control	21	4.454E-07	9.65753E-07	0.073	1	**0.006**
Cardiac muscle contraction | UC	14	2.073E-05	6.80979E-05	0.137	**0.006**	1
Cell cycle - Caulobacter | CD	37	0.0051546	0.000511109	1	0.570	0.026
Cell cycle - Caulobacter | Control	21	0.0052552	0.000526955	0.570	1	**0.013**
Cell cycle - Caulobacter | UC	14	0.0046106	0.000847181	0.026	**0.013**	1
Cell motility and secretion | CD	37	0.0015631	0.000431192	1	0.143	0.044
Cell motility and secretion | Control	21	0.001379	0.000289866	0.143	1	**0.003**
Cell motility and secretion | UC	14	0.0017756	0.0003556	0.044	**0.003**	1
Chloroalkane and chloroalkene degradation | CD	37	0.0017747	0.000453229	1	0.169	0.104
Chloroalkane and chloroalkene degradation | Control	21	0.0019404	0.000206628	0.169	1	**0.010**
Chloroalkane and chloroalkene degradation | UC	14	0.0016031	0.000397071	0.104	**0.010**	1
Citrate cycle (TCA cycle) | CD	37	0.0061242	0.001519857	1	0.021	0.481
Citrate cycle (TCA cycle) | Control	21	0.0052393	0.000884478	0.021	1	**0.014**
Citrate cycle (TCA cycle) | UC	14	0.0062151	0.001035972	0.481	**0.014**	1
Cytoskeleton proteins | CD	37	0.0035794	0.000699379	1	0.213	0.089
Cytoskeleton proteins | Control	21	0.0038779	0.000381148	0.213	1	**0.011**
Cytoskeleton proteins | UC	14	0.0030927	0.000928246	0.089	**0.011**	1
D-Arginine and D-ornithine metabolism | CD	37	1.173E-05	1.31197E-05	1	0.355	**0.008**
D-Arginine and D-ornithine metabolism | Control	21	2.012E-05	2.89051E-05	0.355	1	0.093
D-Arginine and D-ornithine metabolism | UC	14	2.86E-05	3.31746E-05	**0.008**	0.093	1
Ethylbenzene degradation | CD	37	0.0003723	0.000229405	1	0.105	0.133
Ethylbenzene degradation | Control	21	0.0002717	0.000108207	0.105	1	**0.008**
Ethylbenzene degradation | UC	14	0.0004097	0.000124349	0.133	**0.008**	1
Geraniol degradation | CD	37	0.0005015	0.000545966	1	0.124	0.072
Geraniol degradation | Control	21	0.000287	0.00018387	0.124	1	**0.004**
Geraniol degradation | UC	14	0.0006545	0.000479242	0.072	**0.004**	1
Germination | CD	37	0.0003199	0.000173831	1	0.039	0.368
Germination | Control	21	0.0004221	0.00014848	0.039	1	**0.014**
Germination | UC	14	0.00027	0.000171543	0.368	**0.014**	1
Glutathione metabolism | CD	37	0.0019542	0.000597889	1	0.135	0.164
Glutathione metabolism | Control	21	0.0016596	0.000323121	0.135	1	**0.014**

Glutathione metabolism | UC	14	0.0022318	0.000820427	0.164	**0.014**	1
Glycerolipid metabolism | CD	37	0.0039761	0.00068576	1	0.060	0.289
Glycerolipid metabolism | Control	21	0.0043192	0.000517643	0.060	1	**0.014**
Glycerolipid metabolism | UC	14	0.0038719	0.000451198	0.289	**0.014**	1
Huntington's disease | CD	37	0.0002586	0.000145614	1	0.282	0.083
Huntington's disease | Control	21	0.0002114	0.000114383	0.282	1	**0.015**
Huntington's disease | UC	14	0.0003471	0.000178047	0.083	**0.015**	1
Indole alkaloid biosynthesis | CD	37	5.182E-08	2.04564E-07	1	**0.013**	0.498
Indole alkaloid biosynthesis | Control	21	2.421E-09	4.64721E-09	**0.013**	1	0.176
Indole alkaloid biosynthesis | UC	14	1.61E-08	2.65844E-08	0.498	0.176	1
Inorganic ion transport and metabolism | CD	37	0.0021005	0.000631326	1	0.541	**0.013**
Inorganic ion transport and metabolism | Control	21	0.0019199	0.000430234	0.541	1	**0.006**
Inorganic ion transport and metabolism | UC	14	0.0027374	0.000982529	**0.013**	**0.006**	1
Limonene and pinene degradation | CD	37	0.0008529	0.000523819	1	0.085	0.174
Limonene and pinene degradation | Control	21	0.0006507	0.000170867	0.085	1	**0.009**
Limonene and pinene degradation | UC	14	0.0009403	0.000389251	0.174	**0.009**	1
Lipopolysaccharide biosynthesis proteins | CD	37	0.0037177	0.001955587	1	0.072	0.120
Lipopolysaccharide biosynthesis proteins | Control	21	0.0027344	0.001329085	0.072	1	**0.005**
Lipopolysaccharide biosynthesis proteins | UC	14	0.0045324	0.001832543	0.120	**0.005**	1
Lipopolysaccharide biosynthesis | CD	37	0.0021614	0.001527777	1	0.083	0.115
Lipopolysaccharide biosynthesis | Control	21	0.001419	0.000945162	0.083	1	**0.005**
Lipopolysaccharide biosynthesis | UC	14	0.0027856	0.001426587	0.115	**0.005**	1
Lysine biosynthesis | CD	37	0.0084458	0.001117683	1	0.169	0.063
Lysine biosynthesis | Control	21	0.0089649	0.000654788	0.169	1	**0.005**
Lysine biosynthesis | UC	14	0.0078201	0.001326317	0.063	**0.005**	1
Meiosis - yeast | CD	37	1.455E-05	2.43776E-05	1	**0.011**	0.398
Meiosis - yeast | Control	21	3.949E-06	8.21E-06	**0.011**	1	0.210
Meiosis - yeast | UC	14	9.033E-06	1.29885E-05	0.398	0.210	1
Metabolism of cofactors and vitamins | CD	37	0.0011221	0.000226258	1	0.065	0.275
Metabolism of cofactors and vitamins | Control	21	0.0010189	0.000188495	0.065	1	**0.014**
Metabolism of cofactors and vitamins | UC	14	0.0013013	0.000435977	0.275	**0.014**	1
Methane metabolism | CD	37	0.0127053	0.001527261	1	0.139	0.088
Methane metabolism | Control	21	0.0134095	0.00073741	0.139	1	**0.006**
Methane metabolism | UC	14	0.0118841	0.001557333	0.088	**0.006**	1

Nitrogen metabolism | CD	37	0.0074153	0.00072122	1	0.161	0.051
Nitrogen metabolism | Control	21	0.0072462	0.000591253	0.161	1	**0.004**
Nitrogen metabolism | UC	14	0.0078664	0.000726321	0.051	**0.004**	1
Nucleotide excision repair | CD	37	0.0036554	0.000352585	1	0.400	0.042
Nucleotide excision repair | Control	21	0.0037564	0.000259818	0.400	1	**0.012**
Nucleotide excision repair | UC	14	0.0033458	0.000597082	0.042	**0.012**	1
Other ion-coupled transporters | CD	37	0.0144839	0.001648221	1	0.021	0.254
Other ion-coupled transporters | Control	21	0.0137698	0.000998092	0.021	1	**0.004**
Other ion-coupled transporters | UC	14	0.0152736	0.001767603	0.254	**0.004**	1
Parkinson's disease | CD	37	4.008E-06	1.07761E-05	1	0.096	0.148
Parkinson's disease | Control	21	5.404E-07	1.08129E-06	0.096	1	**0.008**
Parkinson's disease | UC	14	2.082E-05	6.80766E-05	0.148	**0.008**	1
Pentose phosphate pathway | CD	37	0.008755	0.000735363	1	**0.010**	0.944
Pentose phosphate pathway | Control	21	0.0092871	0.000732427	**0.010**	1	0.048
Pentose phosphate pathway | UC	14	0.0087373	0.001042016	0.944	0.048	1
Porphyrin and chlorophyll metabolism | CD	37	0.0099987	0.002047134	1	0.109	0.112
Porphyrin and chlorophyll metabolism | Control	21	0.0108852	0.001182942	0.109	1	**0.007**
Porphyrin and chlorophyll metabolism | UC	14	0.0092702	0.001696758	0.112	**0.007**	1
Primary bile acid biosynthesis | CD	37	0.0003545	0.000109218	1	**0.004**	0.948
Primary bile acid biosynthesis | Control	21	0.0004375	8.66092E-05	**0.004**	1	0.019
Primary bile acid biosynthesis | UC	14	0.0003228	0.000145592	0.948	0.019	1
Propanoate metabolism | CD	37	0.0053727	0.000682842	1	**0.015**	0.839
Propanoate metabolism | Control	21	0.0050601	0.000189007	**0.015**	1	0.034
Propanoate metabolism | UC	14	0.0053931	0.000548679	0.839	0.034	1
Protein folding and associated processing | CD	37	0.0059348	0.000917873	1	0.074	0.213
Protein folding and associated processing | Control	21	0.0055151	0.000816966	0.074	1	**0.011**
Protein folding and associated processing | UC	14	0.0063716	0.000909433	0.213	**0.011**	1
RNA transport | CD	37	0.0014904	0.000325551	1	0.445	0.033
RNA transport | Control	21	0.0015656	0.00028152	0.445	1	**0.011**
RNA transport | UC	14	0.0012797	0.000278754	0.033	**0.011**	1
Secondary bile acid biosynthesis | CD	37	0.0003438	0.000111145	1	**0.012**	0.545
Secondary bile acid biosynthesis | Control	21	0.0004224	9.2123E-05	**0.012**	1	**0.011**
Secondary bile acid biosynthesis | UC	14	0.0003034	0.000137932	0.545	**0.011**	1
Sporulation | CD	37	0.0071931	0.003645306	1	0.049	0.113
Sporulation | Control	21	0.009204	0.002310979	0.049	1	**0.003**
Sporulation | UC	14	0.0058355	0.002900545	0.113	**0.003**	1
Staphylococcus aureus infection | CD	37	0.0001509	0.000383921	1	**0.014**	0.393

Staphylococcus aureus infection | Control	21	2.89E-05	1.93432E-05	**0.014**	1	0.244
Staphylococcus aureus infection | UC	14	8.451E-05	0.000117016	0.393	0.244	1
Starch and sucrose metabolism | CD	37	0.0107073	0.001355171	1	0.135	0.066
Starch and sucrose metabolism | Control	21	0.0111916	0.00077383	0.135	1	**0.004**
Starch and sucrose metabolism | UC	14	0.0100096	0.001269198	0.066	**0.004**	1
Thiamine metabolism | CD	37	0.0051367	0.000398252	1	0.537	**0.014**
Thiamine metabolism | Control	21	0.0051023	0.000420329	0.537	1	0.080
Thiamine metabolism | UC	14	0.004747	0.00062333	**0.014**	0.080	1
Toluene degradation | CD	37	0.0013271	0.000835976	1	**0.001**	0.411
Toluene degradation | Control	21	0.0008285	0.00028639	**0.001**	1	**0.001**
Toluene degradation | UC	14	0.001367	0.000579628	0.411	**0.001**	1
Ubiquinone and other terpenoid-quinone biosynthesis | CD	37	0.0018511	0.00077484	1	0.079	0.145
Ubiquinone and other terpenoid-quinone biosynthesis | Control	21	0.0014341	0.000507665	0.079	1	**0.007**
Ubiquinone and other terpenoid-quinone biosynthesis | UC	14	0.002255	0.000950967	0.145	**0.007**	1
Ubiquitin system | CD	37	0.0001106	8.18051E-05	1	0.107	0.136
Ubiquitin system | Control	21	7.745E-05	7.1588E-05	0.107	1	**0.009**
Ubiquitin system | UC	14	0.0001399	7.41685E-05	0.136	**0.009**	1
Valine leucine and isoleucine degradation | CD	37	0.0020982	0.000857031	1	0.134	0.067
Valine leucine and isoleucine degradation | Control	21	0.0017244	0.000334383	0.134	1	**0.004**
Valine leucine and isoleucine degradation | UC	14	0.0023179	0.000635089	0.067	**0.004**	1
Xylene degradation | CD	37	0.000511	0.000168086	1	**0.005**	0.378
Xylene degradation | Control	21	0.0006531	0.000153964	**0.005**	1	0.158
Xylene degradation | UC	14	0.0005704	0.000277117	0.378	0.158	1

Appendix XXII: Microbial metabolic pathways with significant differential abundance in the TI communities of new onset pediatric IBD patients. Relative abundance of KEGG pathways were predicted by PICRUST approach [244]. Statistical significance was assessed using a Kruskal-Wallis test with a Dunn's post hoc test. Values in bold indicate $p <$ Bonferroni corrected significance level of 0.0167. Observations indicate the number of samples from each group employed in the analysis. Mean and std. deviation indicate the mean and std. deviation of the 16S reads obtained. *p*| CD, *p*| Control and *p*| UC indicate the *p* values obtained by comparison to the CD, control and UC, respectively.

Pathway|Variable	Observations	Mean	Std. deviation	*p*|CD	*p*| Control	*p*|UC
Bacterial chemotaxis | CD	27	0.00259863	0.00126955	1	**0.007**	0.759
Bacterial chemotaxis | Control	15	0.00345073	0.00095886	**0.007**	1	0.056
Bacterial chemotaxis | UC	11	0.00261982	0.00102453	0.759	0.056	1
Biosynthesis of ansamycins | CD	27	0.00099511	0.00018352	1	**0.004**	0.343
Biosynthesis of ansamycins | Control	15	0.00124033	0.00027642	**0.004**	1	0.137
Biosynthesis of ansamycins | UC	11	0.00108700	0.00030730	0.343	0.137	1
Chaperones and folding catalysts | CD	27	0.01042170	0.00068944	1	**0.009**	0.600
Chaperones and folding catalysts | Control	15	0.00977513	0.00065704	**0.009**	1	0.102
Chaperones and folding catalysts | UC	11	0.01029909	0.00107174	0.600	0.102	1
Flavone and flavonol biosynthesis | CD	27	0.00011913	0.00009092	1	0.160	**0.012**
Flavone and flavonol biosynthesis | Control	15	0.00008624	0.00008427	0.160	1	0.263
Flavone and flavonol biosynthesis | UC	11	0.00005103	0.00006192	**0.012**	0.263	1
Fructose and mannose metabolism | CD	27	0.01018100	0.00163099	1	**0.014**	0.384
Fructose and mannose metabolism | Control	15	0.01159660	0.00168651	**0.014**	1	0.229
Fructose and mannose metabolism | UC	11	0.01077673	0.00144426	0.384	0.229	1
Lipid biosynthesis proteins | CD	27	0.00596763	0.00048091	1	**0.015**	0.567
Lipid biosynthesis proteins | Control	15	0.00561773	0.00032256	**0.015**	1	0.146
Lipid biosynthesis proteins | UC	11	0.00583282	0.00049337	0.567	0.146	1
Pathways in cancer | CD	27	0.00049996	0.00009556	1	0.020	**0.013**
Pathways in cancer | Control	15	0.00044840	0.00005303	0.020	1	0.734
Pathways in cancer | UC	11	0.00044936	0.00011287	**0.013**	0.734	1
Pentose phosphate pathway | CD	27	0.00845230	0.00077741	1	**0.007**	0.291
Pentose phosphate pathway | Control	15	0.00938300	0.00096211	**0.007**	1	0.214

Pentose phosphate pathway | UC	11	0.00894555	0.00111684	0.291	0.214	1
Taurine and hypotaurine metabolism | CD	27	0.00111233	0.00014473	1	**0.017**	0.733
Taurine and hypotaurine metabolism | Control	15	0.00100427	0.00008095	**0.017**	1	0.024
Taurine and hypotaurine metabolism | UC	11	0.00114009	0.00018515	0.733	0.024	1
Toluene degradation | CD	27	0.00140756	0.00087337	1	**0.008**	0.073
Toluene degradation | Control	15	0.00102580	0.00069048	**0.008**	1	0.582
Toluene degradation | UC	11	0.00101027	0.00040637	0.073	0.582	1
Transporters | CD	27	0.07050419	0.01130273	1	**0.013**	0.750
Transporters | Control	15	0.07906013	0.00700717	**0.013**	1	0.085
Transporters | UC	11	0.07134455	0.01210771	0.750	0.085	1
Tropane piperidine and pyridine alkaloid biosynthesis | CD	27	0.00115156	0.00016940	1	**0.012**	0.306
Tropane piperidine and pyridine alkaloid biosynthesis | Control	15	0.00104567	0.00008743	**0.012**	1	0.266
Tropane piperidine and pyridine alkaloid biosynthesis | UC	11	0.00107855	0.00012637	0.306	0.266	1

Appendix XXIII: Taxa that vary significantly in abundance in Il-10-/- mice in response to *A. parvulum* colonization and/or bismuth administration. Kruskal-Wallis followed by Bonferroni correction for multiple comparisons. Values in bold indicate $p <$ (Bonferroni corrected significance level of 0.0167). Observations indicate the number of samples from each group employed in the analysis. Mean and std. deviation indicate the mean and std. deviation of the 16S reads obtained. *p*| Atopo, *p*| Atopo-Bis, *p*| Bis and p| SPF indicate the *p* values obtained by comparison to the mice treated with *A. parvulum*, *A. parvulum* with bismuth, bismuth and untreated SPF mice, respectively.

Variable	Observations	Mean	Std. deviation	p | Atopo	p | Atopo-Bis	p | Bis	p | SPF
PHYLUM							
Basidiomycota | Atopo	8	0	0	1	1.000	1.000	**0.000**
Basidiomycota | AtopoBis	8	0	0	1.000	1	1.000	**0.000**
Basidiomycota | Bis	8	0	0	1.000	1.000	1	**0.000**
Basidiomycota | SPF	7	5.09E-05	6.42E-05	**0.000**	**0.000**	**0.000**	1
Firmicutes | Atopo	8	0.15954	0.027085	1	0.869	**0.002**	0.194
Firmicutes | AtopoBis	8	0.209969	0.12023	0.869	1	**0.003**	0.255
Firmicutes | Bis	8	0.350887	0.081578	**0.002**	**0.003**	1	0.084
Firmicutes | SPF	7	0.24329	0.115041	0.194	0.255	0.084	1
Cyanobacteria | Atopo	8	0.000112	0.000107	1	0.023	**0.007**	**0.006**
Cyanobacteria | AtopoBis	8	0.000432	0.000314	0.023	1	0.680	0.563
Cyanobacteria | Bis	8	0.000522	0.000389	**0.007**	0.680	1	0.857
Cyanobacteria | SPF	7	0.00074	0.000652	**0.006**	0.563	0.857	1
Fusobacteria | Atopo	8	0.000007	1.37E-05	1	**0.008**	0.089	**0.007**
Fusobacteria | AtopoBis	8	0.01334	0.024208	**0.008**	1	0.341	0.889
Fusobacteria | Bis	8	0.00089	0.002218	0.089	0.341	1	0.289
Fusobacteria | SPF	7	7.71E-05	6.38E-05	**0.007**	0.889	0.289	1
Bacteroidetes | Atopo	8	0.217483	0.0283	1	0.216	**< 0.0001**	0.011
Bacteroidetes | AtopoBis	8	0.129459	0.107836	0.216	1	**0.004**	0.175
Bacteroidetes | Bis	8	0.000419	0.000269	**< 0.0001**	**0.004**	1	0.151
Bacteroidetes | SPF	7	0.057268	0.096359	0.011	0.175	0.151	1
CLASS							
Fusobacteria (class) | Atopo	8	0.000007	1.37E-05	1	**0.008**	0.089	**0.007**
Fusobacteria (class) | AtopoBis	8	0.01334	0.024208	**0.008**	1	0.341	0.889
Fusobacteria (class) | Bis	8	0.00089	0.002218	0.089	0.341	1	0.289
Fusobacteria (class) | SPF	7	7.71E-05	6.38E-05	**0.007**	0.889	0.289	1
Erysipelotrichi | Atopo	8	0.008771	0.004431	1	0.409	0.011	**0.004**
Erysipelotrichi | AtopoBis	8	0.006641	0.005561	0.409	1	0.088	0.037
Erysipelotrichi | Bis	8	0.003558	0.006887	0.011	0.088	1	0.660
Erysipelotrichi | SPF	7	0.001724	0.002967	**0.004**	0.037	0.660	1

Negativicutes | Atopo	8	0.001567	0.00188	1	0.741	0.078	0.198
Negativicutes | AtopoBis	8	3.28E-05	1.71E-05	0.741	1	0.036	0.333
Negativicutes | Bis	8	7.25E-06	6.67E-06	0.078	0.036	1	**0.003**
Negativicutes | SPF	7	0.004445	0.005831	0.198	0.333	**0.003**	1
Clostridia | Atopo	8	0.144006	0.030834	1	0.826	**0.002**	0.111
Clostridia | AtopoBis	8	0.176832	0.095043	0.826	1	**0.005**	0.167
Clostridia | Bis	8	0.323491	0.07562	**0.002**	**0.005**	1	0.184
Clostridia | SPF	7	0.232677	0.115289	0.111	0.167	0.184	1
Agaricomycetes | Atopo	8	0	0	1	1.000	1.000	**0.000**
Agaricomycetes | AtopoBis	8	0	0	1.000	1	1.000	**0.000**
Agaricomycetes | Bis	8	0	0	1.000	1.000	1	**0.000**
Agaricomycetes | SPF	7	5.09E-05	6.42E-05	**0.000**	**0.000**	**0.000**	1
Bacteroidia | Atopo	8	0.217483	0.0283	1	0.216	**< 0.0001**	0.011
Bacteroidia | AtopoBis	8	0.129459	0.107836	0.216	1	**0.004**	0.175
Bacteroidia | Bis	8	0.000419	0.000269	**< 0.0001**	**0.004**	1	0.151
Bacteroidia | SPF	7	0.057268	0.096359	0.011	0.175	0.151	1
Betaproteobacteria | Atopo	8	8.88E-05	0.000203	1	0.067	0.837	**0.001**
Betaproteobacteria | AtopoBis	8	6.08E-05	3.13E-05	0.067	1	0.105	0.102
Betaproteobacteria | Bis	8	2.73E-05	1.66E-05	0.837	0.105	1	**0.001**
Betaproteobacteria | SPF	7	0.000124	6.76E-05	**0.001**	0.102	**0.001**	1
ORDER							
Clostridiales | Atopo	8	0.144006	0.030834	1	0.826	**0.002**	0.111
Clostridiales | AtopoBis	8	0.176796	0.095012	0.826	1	**0.005**	0.167
Clostridiales | Bis	8	0.323491	0.075621	**0.002**	**0.005**	1	0.184
Clostridiales | SPF	7	0.232676	0.115289	0.111	0.167	0.184	1
Alteromonadales | Atopo	8	2.5E-07	7.07E-07	1	0.724	0.282	**0.000**
Alteromonadales | AtopoBis	8	5E-07	9.26E-07	0.724	1	0.470	**0.001**
Alteromonadales | Bis	8	4.98E-05	9.59E-05	0.282	0.470	1	**0.007**
Alteromonadales | SPF	7	0.000528	0.000691	**0.000**	**0.001**	**0.007**	1
Bacteroidales | Atopo	8	0.217483	0.0283	1	0.216	**< 0.0001**	0.011
Bacteroidales | AtopoBis	8	0.129459	0.107836	0.216	1	**0.004**	0.175
Bacteroidales | Bis	8	0.000419	0.000269	**< 0.0001**	**0.004**	1	0.151
Bacteroidales | SPF	7	0.057268	0.096359	0.011	0.175	0.151	1
Oceanospirillales | Atopo	8	0.000004	3.7E-06	1	0.393	0.576	**0.004**
Oceanospirillales | AtopoBis	8	4.63E-05	5.58E-05	0.393	1	0.158	0.037
Oceanospirillales | Bis	8	2.25E-05	4.33E-05	0.576	0.158	1	**0.001**
Oceanospirillales | SPF	7	0.002783	0.003541	**0.004**	0.037	**0.001**	1
Agaricales | Atopo	8	0	0	1	1.000	1.000	**0.000**
Agaricales | AtopoBis	8	0	0	1.000	1	1.000	**0.000**
Agaricales | Bis	8	0	0	1.000	1.000	1	**0.000**
Agaricales | SPF	7	5.09E-05	6.42E-05	**0.000**	**0.000**	**0.000**	1
Actinomycetales | Atopo	8	0.000001	1.07E-06	1	0.446	0.796	**0.002**

Actinomycetales | AtopoBis	8	2.5E-06	2.98E-06	0.446	1	0.615	0.020
Actinomycetales | Bis	8	3.75E-06	7.67E-06	0.796	0.615	1	**0.005**
Actinomycetales | SPF	7	2.46E-05	4.32E-05	**0.002**	0.020	**0.005**	1
Erysipelotrichales | Atopo	8	0.008771	0.004431	1	0.409	0.011	**0.004**
Erysipelotrichales | AtopoBis	8	0.006641	0.005561	0.409	1	0.088	0.037
Erysipelotrichales | Bis	8	0.003558	0.006887	0.011	0.088	1	0.660
Erysipelotrichales | SPF	7	0.001724	0.002967	**0.004**	0.037	0.660	1
Neisseriales | Atopo	8	0	0	1	**0.004**	0.585	0.441
Neisseriales | AtopoBis	8	1.5E-06	1.41E-06	**0.004**	1	0.022	0.048
Neisseriales | Bis	8	2.5E-07	7.07E-07	0.585	0.022	1	0.809
Neisseriales | SPF	7	2.29E-06	6.05E-06	0.441	0.048	0.809	1
Pasteurellales | Atopo	8	2.5E-07	7.07E-07	1	0.594	0.530	**0.000**
Pasteurellales | AtopoBis	8	1.5E-06	2.98E-06	0.594	1	0.925	**0.001**
Pasteurellales | Bis	8	0.000004	9.8E-06	0.530	0.925	1	**0.001**
Pasteurellales | SPF	7	4.03E-05	4.37E-05	**0.000**	**0.001**	**0.001**	1
Chromatiales | Atopo	8	0	0	1	1.000	0.560	**0.000**
Chromatiales | AtopoBis	8	0	0	1.000	1	0.560	**0.000**
Chromatiales | Bis	8	5E-07	1.41E-06	0.560	0.560	1	**0.001**
Chromatiales | SPF	7	9.14E-06	1.05E-05	**0.000**	**0.000**	**0.001**	1
Vibrionales | Atopo	8	4.75E-06	9.07E-06	1	0.191	0.374	**< 0.0001**
Vibrionales | AtopoBis	8	0.00001	1.03E-05	0.191	1	0.677	**0.006**
Vibrionales | Bis	8	1.23E-05	1.62E-05	0.374	0.677	1	**0.002**
Vibrionales | SPF	7	0.030702	0.03792	**< 0.0001**	**0.006**	**0.002**	1
Burkholderiales | Atopo	8	8.78E-05	0.000203	1	0.052	0.978	**0.001**
Burkholderiales | AtopoBis	8	5.55E-05	3.07E-05	0.052	1	0.056	0.137
Burkholderiales | Bis	8	2.23E-05	1.07E-05	0.978	0.056	1	**0.001**
Burkholderiales | SPF	7	0.000117	7.03E-05	**0.001**	0.137	**0.001**	1
Fusobacteriales | Atopo	8	0.000007	1.37E-05	1	**0.008**	0.089	**0.007**
Fusobacteriales | AtopoBis	8	0.01334	0.024208	**0.008**	1	0.341	0.889
Fusobacteriales | Bis	8	0.00089	0.002218	0.089	0.341	1	0.289
Fusobacteriales | SPF	7	7.71E-05	6.38E-05	**0.007**	0.889	0.289	1
Bacillales | Atopo	8	0.00031	0.000196	1	**0.000**	**0.007**	0.115
Bacillales | AtopoBis	8	0.017732	0.019285	**0.000**	1	0.336	0.050
Bacillales | Bis	8	0.011367	0.009925	**0.007**	0.336	1	0.304
Bacillales | SPF	7	0.00321	0.005591	0.115	0.050	0.304	1
Selenomonadales | Atopo	8	0.001567	0.00188	1	0.741	0.078	0.198
Selenomonadales | AtopoBis	8	3.28E-05	1.71E-05	0.741	1	0.036	0.333
Selenomonadales | Bis	8	7.25E-06	6.67E-06	0.078	0.036	1	**0.003**
Selenomonadales | SPF	7	0.004445	0.005831	0.198	0.333	**0.003**	1
Lactobacillales | Atopo	8	0.004886	0.004794	1	0.783	0.296	0.032
Lactobacillales | AtopoBis	8	0.008731	0.012945	0.783	1	0.187	0.060
Lactobacillales | Bis	8	0.012464	0.014999	0.296	0.187	1	**0.002**
Lactobacillales | SPF	7	0.000888	0.00122	0.032	0.060	**0.002**	1
FAMILY							

Staphylococcaceae | Atopo	8	2.75E-06	6.23E-06	1	**0.002**	**0.004**	0.010
Staphylococcaceae | AtopoBis	8	6.53E-05	5.64E-05	**0.002**	1	0.793	0.637
Staphylococcaceae | Bis	8	0.00008	9.17E-05	**0.004**	0.793	1	0.827
Staphylococcaceae | SPF	7	0.000382	0.000666	0.010	0.637	0.827	1
Enterococcaceae | Atopo	8	4.15E-05	1.92E-05	1	**0.000**	0.014	0.040
Enterococcaceae | AtopoBis	8	0.00145	0.001098	**0.000**	1	0.178	0.106
Enterococcaceae | Bis	8	0.001134	0.001063	0.014	0.178	1	0.753
Enterococcaceae | SPF	7	0.000523	0.000991	0.040	0.106	0.753	1
Eubacteriaceae | Atopo	8	1.78E-05	2.77E-05	1	0.061	0.518	**0.001**
Eubacteriaceae | AtopoBis	8	0.000037	1.76E-05	0.061	1	0.220	0.111
Eubacteriaceae | Bis	8	0.000627	0.001677	0.518	0.220	1	**0.005**
Eubacteriaceae | SPF	7	0.00097	0.000799	**0.001**	0.111	**0.005**	1
Ferrimonadaceae | Atopo	8	0	0	1	0.625	1.000	**< 0.0001**
Ferrimonadaceae | AtopoBis	8	2.5E-07	7.07E-07	0.625	1	0.625	**0.000**
Ferrimonadaceae | Bis	8	0	0	1.000	0.625	1	**< 0.0001**
Ferrimonadaceae | SPF	7	0.000281	0.000455	**< 0.0001**	**0.000**	**< 0.0001**	1
Alcanivoracaceae | Atopo	8	0.000003	3.55E-06	1	0.477	0.314	0.064
Alcanivoracaceae | AtopoBis	8	4.43E-05	5.62E-05	0.477	1	0.086	0.245
Alcanivoracaceae | Bis	8	0.000001	2.83E-06	0.314	0.086	1	**0.005**
Alcanivoracaceae | SPF	7	0.000287	0.000417	0.064	0.245	**0.005**	1
Bacteroidaceae | Atopo	8	0.217415	0.028272	1	0.226	**< 0.0001**	**0.007**
Bacteroidaceae | AtopoBis	8	0.129252	0.107884	0.226	1	**0.005**	0.125
Bacteroidaceae | Bis	8	0.000344	0.000287	**< 0.0001**	**0.005**	1	0.239
Bacteroidaceae | SPF	7	0.056777	0.096549	**0.007**	0.125	0.239	1
Oceanospirillaceae | Atopo	8	5E-07	9.26E-07	1	0.424	0.722	**0.000**
Oceanospirillaceae | AtopoBis	8	0.000002	2.83E-06	0.424	1	0.657	**0.003**
Oceanospirillaceae | Bis	8	6.5E-06	1.25E-05	0.722	0.657	1	**0.001**
Oceanospirillaceae | SPF	7	0.002113	0.002626	**0.000**	**0.003**	**0.001**	1
Halomonadaceae | Atopo	8	0	0	1	1.000	1.000	**0.000**
Halomonadaceae | AtopoBis	8	0	0	1.000	1	1.000	**0.000**
Halomonadaceae | Bis	8	0	0	1.000	1.000	1	**0.000**
Halomonadaceae | SPF	7	1.17E-05	1.4E-05	**0.000**	**0.000**	**0.000**	1
Lactobacillaceae | Atopo	8	0.004814	0.004771	1	**0.004**	**0.000**	**0.001**
Lactobacillaceae | AtopoBis	8	2.48E-05	1.85E-05	**0.004**	1	0.440	0.586
Lactobacillaceae | Bis	8	0.000266	0.000722	**0.000**	0.440	1	0.840
Lactobacillaceae | SPF	7	3.91E-05	6.03E-05	**0.001**	0.586	0.840	1
Neisseriaceae | Atopo	8	0	0	1	**0.004**	0.585	0.441
Neisseriaceae | AtopoBis	8	1.5E-06	1.41E-06	**0.004**	1	0.022	0.048
Neisseriaceae | Bis	8	2.5E-07	7.07E-07	0.585	0.022	1	0.809
Neisseriaceae | SPF	7	2.29E-06	6.05E-06	0.441	0.048	0.809	1
Halothiobacillaceae | Atopo	8	0	0	1	1.000	1.000	**< 0.0001**
Halothiobacillaceae |	8	0	0	1.000	1	1.000	**< 0.0001**

AtopoBis							
Halothiobacillaceae | Bis	8	0	0	1.000	1.000	1	**< 0.0001**
Halothiobacillaceae | SPF	7	9.14E-06	1.05E-05	**< 0.0001**	**< 0.0001**	**< 0.0001**	1
Pasteurellaceae | Atopo	8	2.5E-07	7.07E-07	1	0.594	0.530	**0.000**
Pasteurellaceae | AtopoBis	8	1.5E-06	2.98E-06	0.594	1	0.925	**0.001**
Pasteurellaceae | Bis	8	0.000004	9.8E-06	0.530	0.925	1	**0.001**
Pasteurellaceae | SPF	7	4.03E-05	4.37E-05	**0.000**	**0.001**	**0.001**	1
Erysipelotrichaceae | Atopo	8	0.008771	0.004431	1	0.409	0.011	**0.004**
Erysipelotrichaceae | AtopoBis	8	0.006641	0.005561	0.409	1	0.088	0.037
Erysipelotrichaceae | Bis	8	0.003558	0.006887	0.011	0.088	1	0.660
Erysipelotrichaceae | SPF	7	0.001724	0.002967	**0.004**	0.037	0.660	1
Fusobacteriaceae | Atopo	8	0.000007	1.37E-05	1	**0.008**	0.089	**0.007**
Fusobacteriaceae | AtopoBis	8	0.01334	0.024208	**0.008**	1	0.341	0.889
Fusobacteriaceae | Bis	8	0.00089	0.002218	0.089	0.341	1	0.289
Fusobacteriaceae | SPF	7	7.71E-05	6.38E-05	**0.007**	0.889	0.289	1
Bacillaceae | Atopo	8	0.000296	0.000189	1	**0.001**	0.024	0.311
Bacillaceae | AtopoBis	8	0.016992	0.018066	**0.001**	1	0.284	0.028
Bacillaceae | Bis	8	0.011192	0.009815	0.024	0.284	1	0.244
Bacillaceae | SPF	7	0.002803	0.005694	0.311	0.028	0.244	1
Listeriaceae | Atopo	8	7.5E-07	1.04E-06	1	**0.006**	0.029	0.477
Listeriaceae | AtopoBis	8	1.45E-05	1.3E-05	**0.006**	1	0.590	0.055
Listeriaceae | Bis	8	0.00001	9.5E-06	0.029	0.590	1	0.161
Listeriaceae | SPF	7	9.14E-06	2.16E-05	0.477	0.055	0.161	1
Streptococcaceae | Atopo	8	0.000015	1.5E-05	1	0.182	**0.001**	0.056
Streptococcaceae | AtopoBis	8	0.006866	0.012053	0.182	1	0.046	0.532
Streptococcaceae | Bis	8	0.010793	0.01558	**0.001**	0.046	1	0.193
Streptococcaceae | SPF	7	0.000169	0.000201	0.056	0.532	0.193	1
Vibrionaceae | Atopo	8	4.75E-06	9.07E-06	1	0.191	0.374	**< 0.0001**
Vibrionaceae | AtopoBis	8	0.00001	1.03E-05	0.191	1	0.677	**0.006**
Vibrionaceae | Bis	8	1.23E-05	1.62E-05	0.374	0.677	1	**0.002**
Vibrionaceae | SPF	7	0.030702	0.03792	**< 0.0001**	**0.006**	**0.002**	1
Methylococcaceae | Atopo	8	0	0	1	1.000	1.000	0.027
Methylococcaceae | AtopoBis	8	0	0	1.000	1	1.000	0.027
Methylococcaceae | Bis	8	0	0	1.000	1.000	1	0.027
Methylococcaceae | SPF	7	8.57E-07	1.57E-06	0.027	0.027	0.027	1
Sutterellaceae | Atopo	8	6.25E-06	1.12E-05	1	0.020	0.708	**0.007**
Sutterellaceae | AtopoBis	8	2.43E-05	2.45E-05	0.020	1	0.050	0.645
Sutterellaceae | Bis	8	0.000005	6.76E-06	0.708	0.050	1	0.019
Sutterellaceae | SPF	7	6.71E-05	8.81E-05	**0.007**	0.645	0.019	1
Veillonellaceae | Atopo	8	0.001562	0.001876	1	0.815	0.085	0.197
Veillonellaceae | AtopoBis	8	2.53E-05	1.54E-05	0.815	1	0.050	0.287
Veillonellaceae | Bis	8	6.75E-06	6.76E-06	0.085	0.050	1	**0.003**

Veillonellaceae | SPF	7	0.004444	0.005831	0.197	0.287	**0.003**	1
Catabacteriaceae | Atopo	8	5E-07	1.41E-06	1	0.643	0.220	**0.004**
Catabacteriaceae | AtopoBis	8	0	0	0.643	1	0.091	**0.001**
Catabacteriaceae | Bis	8	0.000628	0.001185	0.220	0.091	1	0.097
Catabacteriaceae | SPF	7	0.000228	0.000402	**0.004**	**0.001**	0.097	1
Shewanellaceae | Atopo	8	0	0	1	1.000	0.352	**0.000**
Shewanellaceae | AtopoBis	8	0	0	1.000	1	0.352	**0.000**
Shewanellaceae | Bis	8	5E-07	9.26E-07	0.352	0.352	1	**0.003**
Shewanellaceae | SPF	7	4.89E-05	8.53E-05	**0.000**	**0.000**	**0.003**	1
Hahellaceae | Atopo	8	5E-07	9.26E-07	1	0.427	0.947	**0.002**
Hahellaceae | AtopoBis	8	0	0	0.427	1	0.389	**0.000**
Hahellaceae | Bis	8	1.25E-06	2.82E-06	0.947	0.389	1	**0.003**
Hahellaceae | SPF	7	0.000112	0.00014	**0.002**	**0.000**	**0.003**	1
Actinomycetaceae | Atopo	8	2.5E-07	7.07E-07	1	0.190	0.325	**0.000**
Actinomycetaceae | AtopoBis	8	1.75E-06	2.25E-06	0.190	1	0.743	0.017
Actinomycetaceae | Bis	8	2.75E-06	6.23E-06	0.325	0.743	1	**0.007**
Actinomycetaceae | SPF	7	2.31E-05	4.29E-05	**0.000**	0.017	**0.007**	1
Alteromonadaceae | Atopo	8	2.5E-07	7.07E-07	1	1.000	0.274	**0.001**
Alteromonadaceae | AtopoBis	8	2.5E-07	7.07E-07	1.000	1	0.274	**0.001**
Alteromonadaceae | Bis	8	4.93E-05	9.51E-05	0.274	0.274	1	0.027
Alteromonadaceae | SPF	7	0.000198	0.000214	**0.001**	**0.001**	0.027	1
Lycoperdaceae | Atopo	8	0	0	1	1.000	1.000	**0.000**
Lycoperdaceae | AtopoBis	8	0	0	1.000	1	1.000	**0.000**
Lycoperdaceae | Bis	8	0	0	1.000	1.000	1	**0.000**
Lycoperdaceae | SPF	7	5.09E-05	6.42E-05	**0.000**	**0.000**	**0.000**	1
Peptostreptococcaceae | Atopo	8	0	0	1	0.308	0.352	**0.006**
Peptostreptococcaceae | AtopoBis	8	1.25E-06	2.38E-06	0.308	1	0.929	0.078
Peptostreptococcaceae | Bis	8	0.000001	1.85E-06	0.352	0.929	1	0.065
Peptostreptococcaceae | SPF	7	7.43E-06	8.7E-06	**0.006**	0.078	0.065	1
GENUS							
Morganella | Atopo	8	0	0	1	1.000	1.000	**0.000**
Morganella | AtopoBis	8	0	0	1.000	1	1.000	**0.000**
Morganella | Bis	8	0	0	1.000	1.000	1	**0.000**
Morganella | SPF	7	5.09E-05	6.42E-05	**0.000**	**0.000**	**0.000**	1
Erwinia | Atopo	8	2.55E-05	1.74E-05	1	0.826	0.710	**0.003**
Erwinia | AtopoBis	8	9.93E-05	0.000122	0.826	1	0.880	**0.006**
Erwinia | Bis	8	0.020397	0.039265	0.710	0.880	1	0.010
Erwinia | SPF	7	0.00184	0.001984	**0.003**	**0.006**	0.010	1
Peptostreptococcus | Atopo	8	0	0	1	0.308	0.352	**0.006**
Peptostreptococcus | AtopoBis	8	1.25E-06	2.38E-06	0.308	1	0.929	0.078

Peptostreptococcus | Bis	8	0.000001	1.85E-06	0.352	0.929	1	0.065
Peptostreptococcus | SPF	7	7.43E-06	8.7E-06	**0.006**	0.078	0.065	1
Dorea | Atopo	8	5.25E-06	5.95E-06	1	0.757	0.527	0.041
Dorea | AtopoBis	8	1.15E-05	1.58E-05	0.757	1	0.346	0.081
Dorea | Bis	8	4.75E-06	7.85E-06	0.527	0.346	1	**0.008**
Dorea | SPF	7	2.69E-05	2.02E-05	0.041	0.081	**0.008**	1
Ruminococcus | Atopo	8	0.000125	0.000179	1	0.030	0.021	**0.002**
Ruminococcus | AtopoBis	8	0.000883	0.001221	0.030	1	0.891	0.341
Ruminococcus | Bis	8	0.000797	0.000803	0.021	0.891	1	0.412
Ruminococcus | SPF	7	0.009368	0.016302	**0.002**	0.341	0.412	1
Kangiella | Atopo	8	0	0	1	1.000	0.548	**0.003**
Kangiella | AtopoBis	8	0	0	1.000	1	0.548	**0.003**
Kangiella | Bis	8	2.5E-07	7.07E-07	0.548	0.548	1	0.015
Kangiella | SPF	7	8.86E-06	1.67E-05	**0.003**	**0.003**	0.015	1
Enterovibrio | Atopo	8	0	0	1	1.000	1.000	**< 0.0001**
Enterovibrio | AtopoBis	8	0	0	1.000	1	1.000	**< 0.0001**
Enterovibrio | Bis	8	0	0	1.000	1.000	1	**< 0.0001**
Enterovibrio | SPF	7	0.000176	0.000185	**< 0.0001**	**< 0.0001**	**< 0.0001**	1
Coprobacillus | Atopo	8	0.002304	0.001357	1	**0.003**	**0.008**	**0.004**
Coprobacillus | AtopoBis	8	0.000515	0.001211	**0.003**	1	0.772	0.970
Coprobacillus | Bis	8	0.000401	0.000906	**0.008**	0.772	1	0.750
Coprobacillus | SPF	7	0.000355	0.000698	**0.004**	0.970	0.750	1
Actinomyces | Atopo	8	2.5E-07	7.07E-07	1	0.190	0.325	**0.000**
Actinomyces | AtopoBis	8	1.75E-06	2.25E-06	0.190	1	0.743	0.017
Actinomyces | Bis	8	2.75E-06	6.23E-06	0.325	0.743	1	**0.007**
Actinomyces | SPF	7	2.31E-05	4.29E-05	**0.000**	0.017	**0.007**	1
Marinomonas | Atopo	8	5E-07	9.26E-07	1	0.694	0.694	**0.002**
Marinomonas | AtopoBis	8	7.5E-07	1.04E-06	0.694	1	0.431	**0.007**
Marinomonas | Bis	8	2.5E-07	7.07E-07	0.694	0.431	1	**0.001**
Marinomonas | SPF	7	0.00029	0.000371	**0.002**	**0.007**	**0.001**	1
Vibrio | Atopo	8	4.75E-06	9.07E-06	1	0.216	0.381	**< 0.0001**
Vibrio | AtopoBis	8	9.5E-06	9.37E-06	0.216	1	0.718	**0.006**
Vibrio | Bis	8	1.23E-05	1.62E-05	0.381	0.718	1	**0.002**
Vibrio | SPF	7	0.030436	0.037658	**< 0.0001**	**0.006**	**0.002**	1
Dialister | Atopo	8	2.75E-06	6.23E-06	1	0.020	0.534	0.129
Dialister | AtopoBis	8	1.33E-05	1.33E-05	0.020	1	**0.003**	0.465
Dialister | Bis	8	5E-07	9.26E-07	0.534	**0.003**	1	0.034
Dialister | SPF	7	7.14E-06	6.2E-06	0.129	0.465	0.034	1
Halothiobacillus | Atopo	8	0	0	1	1.000	1.000	**< 0.0001**
Halothiobacillus | AtopoBis	8	0	0	1.000	1	1.000	**< 0.0001**
Halothiobacillus | Bis	8	0	0	1.000	1.000	1	**< 0.0001**
Halothiobacillus | SPF	7	9.14E-06	1.05E-05	**< 0.0001**	**< 0.0001**	**< 0.0001**	1
Trichococcus | Atopo	8	2.25E-06	4.06E-06	1	**0.007**	0.041	0.009

Trichococcus | AtopoBis	8	0	0	**0.007**	1	0.519	1.000
Trichococcus | Bis	8	2.5E-07	7.07E-07	0.041	0.519	1	0.533
Trichococcus | SPF	7	0	0	0.009	1.000	0.533	1
Nitrincola | Atopo	8	0	0	1	0.577	1.000	**0.001**
Nitrincola | AtopoBis	8	7.5E-07	2.12E-06	0.577	1	0.577	**0.003**
Nitrincola | Bis	8	0	0	1.000	0.577	1	**0.001**
Nitrincola | SPF	7	0.00011	0.000168	**0.001**	**0.003**	**0.001**	1
Serratia | Atopo	8	0	0	1	0.198	1.000	**0.000**
Serratia | AtopoBis	8	1.5E-06	2.33E-06	0.198	1	0.198	0.009
Serratia | Bis	8	0	0	1.000	0.198	1	**0.000**
Serratia | SPF	7	0.000385	0.000474	**0.000**	0.009	**0.000**	1
Ferrimonas | Atopo	8	0	0	1	0.625	1.000	**< 0.0001**
Ferrimonas | AtopoBis	8	2.5E-07	7.07E-07	0.625	1	0.625	**0.000**
Ferrimonas | Bis	8	0	0	1.000	0.625	1	**< 0.0001**
Ferrimonas | SPF	7	0.000281	0.000455	**< 0.0001**	**0.000**	**< 0.0001**	1
Butyrivibrio | Atopo	8	2.5E-07	7.07E-07	1	1.000	0.534	0.032
Butyrivibrio | AtopoBis	8	2.5E-07	7.07E-07	1.000	1	0.534	0.032
Butyrivibrio | Bis	8	0	0	0.534	0.534	1	**0.006**
Butyrivibrio | SPF	7	1.14E-06	1.07E-06	0.032	0.032	**0.006**	1
Oscillospira | Atopo	8	1.5E-06	2.33E-06	1	0.055	0.144	**0.000**
Oscillospira | AtopoBis	8	6.25E-06	4.33E-06	0.055	1	0.646	0.083
Oscillospira | Bis	8	0.001431	0.002744	0.144	0.646	1	0.030
Oscillospira | SPF	7	0.001426	0.002879	**0.000**	0.083	0.030	1
Epulopiscium | Atopo	8	3.25E-06	6.14E-06	1	0.434	0.099	0.014
Epulopiscium | AtopoBis	8	7.5E-07	1.49E-06	0.434	1	0.015	**0.001**
Epulopiscium | Bis	8	0.0024	0.005864	0.099	0.015	1	0.393
Epulopiscium | SPF	7	3.49E-05	4.82E-05	0.014	**0.001**	0.393	1
Escherichia | Atopo	8	0.000004	5.01E-06	1	0.029	0.557	0.351
Escherichia | AtopoBis	8	0	0	0.029	1	0.110	**0.002**
Escherichia | Bis	8	2.08E-05	3.96E-05	0.557	0.110	1	0.133
Escherichia | SPF	7	5.43E-06	3.6E-06	0.351	**0.002**	0.133	1
Alkalimonas | Atopo	8	0	0	1	1.000	1.000	**0.006**
Alkalimonas | AtopoBis	8	0	0	1.000	1	1.000	**0.006**
Alkalimonas | Bis	8	0	0	1.000	1.000	1	**0.006**
Alkalimonas | SPF	7	1.54E-05	2.27E-05	**0.006**	**0.006**	**0.006**	1
Listeria | Atopo	8	7.5E-07	1.04E-06	1	**0.006**	0.029	0.477
Listeria | AtopoBis	8	1.45E-05	1.3E-05	**0.006**	1	0.590	0.055
Listeria | Bis	8	0.00001	9.5E-06	0.029	0.590	1	0.161
Listeria | SPF	7	9.14E-06	2.16E-05	0.477	0.055	0.161	1
Streptococcus | Atopo	8	0.000015	1.5E-05	1	0.182	**0.001**	0.056
Streptococcus | AtopoBis	8	0.006866	0.012053	0.182	1	0.046	0.532
Streptococcus | Bis	8	0.010793	0.01558	**0.001**	0.046	1	0.193
Streptococcus | SPF	7	0.000169	0.000201	0.056	0.532	0.193	1
Actinobacillus | Atopo	8	0	0	1	0.681	0.328	**0.000**
Actinobacillus | AtopoBis	8	2.5E-07	7.07E-07	0.681	1	0.571	**0.001**

Actinobacillus | Bis	8	2.25E-06	5.6E-06	0.328	0.571	1	**0.005**
Actinobacillus | SPF	7	0.000024	3E-05	**0.000**	**0.001**	**0.005**	1
Roseburia | Atopo	8	4.63E-05	2.54E-05	1	0.773	0.040	0.321
Roseburia | AtopoBis	8	4.58E-05	3.57E-05	0.773	1	0.078	0.203
Roseburia | Bis	8	1.58E-05	1.49E-05	0.040	0.078	1	**0.003**
Roseburia | SPF	7	0.000131	0.000129	0.321	0.203	**0.003**	1
Bacillus | Atopo	8	0.000291	0.000185	1	**0.001**	0.023	0.298
Bacillus | AtopoBis	8	0.015146	0.015304	**0.001**	1	0.284	0.029
Bacillus | Bis	8	0.009723	0.008202	0.023	0.284	1	0.250
Bacillus | SPF	7	0.002753	0.005611	0.298	0.029	0.250	1
Parabacteroides | Atopo	8	1.78E-05	2.71E-05	1	0.195	0.562	0.036
Parabacteroides | AtopoBis	8	2.55E-05	2.16E-05	0.195	1	0.061	0.397
Parabacteroides | Bis	8	7.5E-06	6.57E-06	0.562	0.061	1	**0.008**
Parabacteroides | SPF	7	4.71E-05	3.33E-05	0.036	0.397	**0.008**	1
Sarcina | Atopo	8	2.5E-07	7.07E-07	1	**0.005**	0.009	0.317
Sarcina | AtopoBis	8	0.000849	0.001397	**0.005**	1	0.815	0.082
Sarcina | Bis	8	0.0014	0.001933	0.009	0.815	1	0.131
Sarcina | SPF	7	4.57E-06	6.29E-06	0.317	0.082	0.131	1
Enterococcus | Atopo	8	4.15E-05	1.92E-05	1	**0.000**	0.016	0.040
Enterococcus | AtopoBis	8	0.001436	0.00108	**0.000**	1	0.161	0.100
Enterococcus | Bis	8	0.001117	0.001056	0.016	0.161	1	0.773
Enterococcus | SPF	7	0.000521	0.000991	0.040	0.100	0.773	1
Carnobacterium | Atopo	8	2.5E-07	7.07E-07	1	**0.006**	0.045	0.543
Carnobacterium | AtopoBis	8	0.000035	3.75E-05	**0.006**	1	0.451	0.040
Carnobacterium | Bis	8	0.000013	2.02E-05	0.045	0.451	1	0.185
Carnobacterium | SPF	7	0.000002	3.46E-06	0.543	0.040	0.185	1
Coprococcus | Atopo	8	0.015587	0.023462	1	**0.001**	0.012	0.220
Coprococcus | AtopoBis	8	8.13E-05	4.53E-05	**0.001**	1	0.433	0.050
Coprococcus | Bis	8	0.001788	0.003606	0.012	0.433	1	0.228
Coprococcus | SPF	7	0.013735	0.024551	0.220	0.050	0.228	1
Enterobacter | Atopo	8	0.000001	2.14E-06	1	0.849	0.333	**0.000**
Enterobacter | AtopoBis	8	7.5E-07	1.04E-06	0.849	1	0.437	**0.001**
Enterobacter | Bis	8	5.75E-06	9.59E-06	0.333	0.437	1	**0.007**
Enterobacter | SPF	7	0.001365	0.001754	**0.000**	**0.001**	**0.007**	1
Neisseria | Atopo	8	0	0	1	**0.003**	0.538	0.387
Neisseria | AtopoBis	8	1.5E-06	1.41E-06	**0.003**	1	0.014	0.032
Neisseria | Bis	8	2.5E-07	7.07E-07	0.538	0.014	1	0.785
Neisseria | SPF	7	1.43E-06	3.78E-06	0.387	0.032	0.785	1
Photobacterium | Atopo	8	0	0	1	0.222	1.000	**< 0.0001**
Photobacterium | AtopoBis	8	5E-07	1.41E-06	0.222	1	0.222	**< 0.0001**
Photobacterium | Bis	8	0	0	1.000	0.222	1	**< 0.0001**
Photobacterium | SPF	7	8.46E-05	9.93E-05	**< 0.0001**	**< 0.0001**	**< 0.0001**	1
Brenneria | Atopo	8	0	0	1	1.000	1.000	**0.000**
Brenneria | AtopoBis	8	0	0	1.000	1	1.000	**0.000**
Brenneria | Bis	8	0	0	1.000	1.000	1	**0.000**

Brenneria | SPF	7	0.000313	0.000453	**0.000**	**0.000**	**0.000**	1
Oceanobacillus | Atopo	8	0	0	1	0.465	**0.005**	1.000
Oceanobacillus | AtopoBis	8	2.5E-07	7.07E-07	0.465	1	0.026	0.480
Oceanobacillus | Bis	8	1.25E-06	1.49E-06	**0.005**	0.026	1	**0.006**
Oceanobacillus | SPF	7	0	0	1.000	0.480	**0.006**	1
Lactobacillus | Atopo	8	0.004814	0.004771	1	**0.000**	**< 0.0001**	**< 0.0001**
Lactobacillus | AtopoBis	8	2.48E-05	1.85E-05	**0.000**	1	0.277	0.441
Lactobacillus | Bis	8	0.000266	0.000722	**< 0.0001**	0.277	1	0.774
Lactobacillus | SPF	7	3.91E-05	6.03E-05	**< 0.0001**	0.441	0.774	1
Xanthomonas | Atopo	8	2.5E-07	7.07E-07	1	0.620	0.138	**0.007**
Xanthomonas | AtopoBis	8	0	0	0.620	1	0.052	**0.002**
Xanthomonas | Bis	8	0.000415	0.00079	0.138	0.052	1	0.163
Xanthomonas | SPF	7	0.000012	2.07E-05	**0.007**	**0.002**	0.163	1
Sutterella | Atopo	8	6.25E-06	1.12E-05	1	0.009	0.657	**0.003**
Sutterella | AtopoBis	8	2.43E-05	2.45E-05	0.009	1	0.026	0.586
Sutterella | Bis	8	0.000005	6.76E-06	0.657	0.026	1	0.009
Sutterella | SPF	7	6.71E-05	8.81E-05	**0.003**	0.586	0.009	1
Staphylococcus | Atopo	8	2.5E-06	5.53E-06	1	**0.000**	**0.001**	**0.003**
Staphylococcus | AtopoBis	8	6.53E-05	5.64E-05	**0.000**	1	0.744	0.558
Staphylococcus | Bis	8	0.00008	9.17E-05	**0.001**	0.744	1	0.785
Staphylococcus | SPF	7	0.000382	0.000666	**0.003**	0.558	0.785	1
Lachnobacterium | Atopo	8	8.25E-06	1.52E-05	1	0.072	0.275	0.068
Lachnobacterium | AtopoBis	8	0	0	0.072	1	**0.006**	**0.001**
Lachnobacterium | Bis	8	0.004256	0.008649	0.275	**0.006**	1	0.418
Lachnobacterium | SPF	7	0.000462	0.000744	0.068	**0.001**	0.418	1
Vagococcus | Atopo	8	0	0	1	**0.001**	**0.004**	0.129
Vagococcus | AtopoBis	8	0.000014	2.48E-05	**0.001**	1	0.570	0.049
Vagococcus | Bis	8	0.000017	2.43E-05	**0.004**	0.570	1	0.144
Vagococcus | SPF	7	1.71E-06	2.43E-06	0.129	0.049	0.144	1
Leclercia | Atopo	8	0.000001	1.51E-06	1	0.298	0.501	**< 0.0001**
Leclercia | AtopoBis	8	0.000001	2.83E-06	0.298	1	0.707	**< 0.0001**
Leclercia | Bis	8	5E-07	9.26E-07	0.501	0.707	1	**< 0.0001**
Leclercia | SPF	7	0.002838	0.003752	**< 0.0001**	**< 0.0001**	**< 0.0001**	1
Fusobacterium | Atopo	8	0.000007	1.37E-05	1	**0.006**	0.065	**0.005**
Fusobacterium | AtopoBis	8	0.01334	0.024208	**0.006**	1	0.307	0.865
Fusobacterium | Bis	8	0.00089	0.002218	0.065	0.307	1	0.249
Fusobacterium | SPF	7	0.000076	6.31E-05	**0.005**	0.865	0.249	1
Citrobacter | Atopo	8	0	0	1	1.000	1.000	**0.000**
Citrobacter | AtopoBis	8	0	0	1.000	1	1.000	**0.000**
Citrobacter | Bis	8	0	0	1.000	1.000	1	**0.000**
Citrobacter | SPF	7	2.37E-05	3.3E-05	**0.000**	**0.000**	**0.000**	1
Hahella | Atopo	8	5E-07	9.26E-07	1	0.265	0.925	**0.000**
Hahella | AtopoBis	8	0	0	0.265	1	0.229	**< 0.0001**

Hahella | Bis	8	1.25E-06	2.82E-06	0.925	0.229	1	**0.000**
Hahella | SPF	7	0.000112	0.00014	**0.000**	**< 0.0001**	**0.000**	1
Alcanivorax | Atopo	8	0.000003	3.55E-06	1	0.481	0.085	0.045
Alcanivorax | AtopoBis	8	4.43E-05	5.62E-05	0.481	1	0.019	0.170
Alcanivorax | Bis	8	0	0	0.085	0.019	1	**0.001**
Alcanivorax | SPF	7	0.000278	0.000402	0.045	0.170	**0.001**	1
Facklamia | Atopo	8	0	0	1	1.000	1.000	**0.003**
Facklamia | AtopoBis	8	0	0	1.000	1	1.000	**0.003**
Facklamia | Bis	8	0	0	1.000	1.000	1	**0.003**
Facklamia | SPF	7	2.31E-05	4.3E-05	**0.003**	**0.003**	**0.003**	1
Faecalibacterium | Atopo	8	7.93E-05	0.000172	1	0.010	0.687	**0.007**
Faecalibacterium | AtopoBis	8	0.000114	7.48E-05	0.010	1	**0.004**	0.811
Faecalibacterium | Bis	8	2.68E-05	2.3E-05	0.687	**0.004**	1	**0.003**
Faecalibacterium | SPF	7	0.000177	0.000154	**0.007**	0.811	**0.003**	1
Eubacterium | Atopo	8	1.78E-05	2.77E-05	1	0.024	0.417	**0.000**
Eubacterium | AtopoBis	8	0.000037	1.76E-05	0.024	1	0.130	0.052
Eubacterium | Bis	8	0.000627	0.001677	0.417	0.130	1	**0.001**
Eubacterium | SPF	7	0.00097	0.000799	**0.000**	0.052	**0.001**	1
Shewanella | Atopo	8	0	0	1	1.000	0.144	**< 0.0001**
Shewanella | AtopoBis	8	0	0	1.000	1	0.144	**< 0.0001**
Shewanella | Bis	8	5E-07	9.26E-07	0.144	0.144	1	**< 0.0001**
Shewanella | SPF	7	4.89E-05	8.53E-05	**< 0.0001**	**< 0.0001**	**< 0.0001**	1
Tatumella | Atopo	8	0	0	1	0.454	1.000	**< 0.0001**
Tatumella | AtopoBis	8	2.5E-07	7.07E-07	0.454	1	0.454	**0.000**
Tatumella | Bis	8	0	0	1.000	0.454	1	**< 0.0001**
Tatumella | SPF	7	5.34E-05	6.73E-05	**< 0.0001**	**0.000**	**< 0.0001**	1
Bacteroides | Atopo	8	0.217415	0.028272	1	0.075	**< 0.0001**	**0.000**
Bacteroides | AtopoBis	8	0.129252	0.107884	0.075	1	**0.000**	0.027
Bacteroides | Bis	8	0.000344	0.000287	**< 0.0001**	**0.000**	1	0.083
Bacteroides | SPF	7	0.056777	0.096549	**0.000**	0.027	0.083	1

Appendix XXIV: List of all differentially expressed proteins and their variable importance in projection scores (VIP) derived from the calculated PLS-DA model.

Variable	Comp 1	Comp 2	Comp 3
	VIP	VIP	VIP
General transcription factor IIA subunit 1;TFIIA p19 subunit;TFIIA p35 subunit;TFIIAL;Transcription initiation factor IIA alpha chain;Transcription initiation factor IIA beta chain;Transcription initiation factor IIA subunit 1;Transcription initiation factor TFIIA 42 kDa subunit	2.513	1.975	1.814
Angiotensin-binding protein;Microsomal endopeptidase;Mitochondrial oligopeptidase M;Neurolysin, mitochondrial;Neurotensin endopeptidase	2.296	1.901	1.745
Defensin, alpha 5;Defensin-5	2.013	1.575	1.399
Mineral dust-induced gene protein;MYC-induced nuclear antigen;Nucleolar protein 52	1.942	1.793	1.585
Glutaminase kidney isoform, mitochondrial;K-glutaminase;L-glutamine amidohydrolase	1.880	1.538	1.358
Ethanolaminephosphotransferase 1;Selenoprotein I;Putative uncharacterized protein ENSP00000385426;Putative uncharacterized protein ENSP00000391804	1.853	1.918	1.702
18S rRNA dimethylase;DIM1 dimethyladenosine transferase 1-like;Probable dimethyladenosine transferase;S-adenosylmethionine-6-N,N-adenosyl(rRNA) dimethyltransferase	1.790	1.524	1.435
6PF-2-K/Fru-2,6-P2ase heart-type isozyme;6-phosphofructo-2-kinase;6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 2;Fructose-2,6-bisphosphatase	1.717	1.442	1.275
Armadillo repeat-containing protein 8;cDNA FLJ56387, highly similar to Mus musculus armadillo repeat containing 8 (Armc8), mRNA;Putative uncharacterized protein ARMC8;Armadillo repeat containing 8, isoform CRA_g;cDNA FLJ53383, highly similar to Homo sapiens armadillo repeat containing 8 (ARMC8), transcript variant 2, mRNA	1.692	1.548	1.376
Aconitase 2, mitochondrial;Aconitate hydratase, mitochondrial;Citrate hydro-lyase;cDNA FLJ60429, highly similar to Aconitate hydratase, mitochondrial (EC 4.2.1.3);cDNA FLJ50886, highly similar to Aconitate hydratase, mitochondrial(EC 4.2.1.3)	1.645	1.338	1.191
2C4D;Class II mMOB1;Mob1 homolog 3;Mps one binder kinase activator-like 3;Preimplantation protein 3;cDNA FLJ52887, highly similar to Preimplantation protein 3	1.634	1.359	1.203
Iron-sulfur subunit of complex II;Succinate dehydrogenase [ubiquinone] iron-sulfur subunit, mitochondrial	1.612	1.266	1.216
Reticulocalbin-1;cDNA FLJ55835, highly similar to Reticulocalbin-1	1.593	1.238	1.093
DRB sensitivity-inducing factor 14 kDa subunit;DRB sensitivity-inducing factor small subunit;Transcription elongation factor SPT4	1.592	1.247	1.102

Rhodanese;Thiosulfate sulfurtransferase	1.590	1.257	1.112
22 kDa protein;CP-22;Sorcin;V19;Putative uncharacterized protein SRI;cDNA FLJ60640, highly similar to Sorcin;cDNA FLJ54267, moderately similar to Sorcin	1.589	1.282	1.144
Flavoprotein subunit of complex II;Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial	1.576	1.225	1.169
Acetylneuraminyl hydrolase;G9 sialidase;Lysosomal sialidase;N-acetyl-alpha-neuraminidase 1;Sialidase-1	1.562	1.955	1.773
Beta-IV spectrin;Spectrin beta chain, brain 3;Spectrin, non-erythroid beta chain 3;Putative uncharacterized protein SPTBN4	1.557	1.291	1.194
Translocation protein SEC63 homolog	1.542	1.555	1.373
Epidermal-type fatty acid-binding protein;Fatty acid-binding protein 5;Fatty acid-binding protein, epidermal;Psoriasis-associated fatty acid-binding protein homolog	1.540	1.206	1.071
Complex I-51kD;NADH dehydrogenase [ubiquinone] flavoprotein 1, mitochondrial;NADH dehydrogenase flavoprotein 1;NADH-ubiquinone oxidoreductase 51 kDa subunit;cDNA FLJ57949, highly similar to NADH-ubiquinone oxidoreductase 51 kDa subunit, mitochondrial (EC 1.6.5.3);cDNA, FLJ79021, highly similar to NADH-ubiquinone oxidoreductase 51 kDa subunit, mitochondrial (EC 1.6.5.3)	1.530	1.205	1.205
Calregulin;Calreticulin;CRP55;Endoplasmic reticulum resident protein 60;grp60;HACBP;cDNA FLJ58668, highly similar to Calreticulin	1.528	1.220	1.077
UDP-glucose 6-dehydrogenase;cDNA FLJ60093, highly similar to UDP-glucose 6-dehydrogenase (EC 1.1.1.22)	1.522	1.183	1.052
4-alpha-glucanotransferase;Amylo-alpha-1,6-glucosidase;Dextrin 6-alpha-D-glucosidase;Glycogen debrancher;Glycogen debranching enzyme;Oligo-1,4-1,4-glucantransferase	1.516	1.200	1.072
Malic enzyme 2;NAD-dependent malic enzyme, mitochondrial	1.514	1.225	1.081
Delta(3),delta(2)-enoyl-CoA isomerase;Diazepam-binding inhibitor-related protein 1;Dodecenoyl-CoA isomerase;DRS-1;Hepatocellular carcinoma-associated antigen 88;Peroxisomal 3,2-trans-enoyl-CoA isomerase;Renal carcinoma antigen NY-REN-1;Putative uncharacterized protein PECI	1.513	1.178	1.046
Complex I-75kD;NADH-ubiquinone oxidoreductase 75 kDa subunit, mitochondrial;cDNA FLJ60586, highly similar to NADH-ubiquinone oxidoreductase 75 kDa subunit, mitochondrial (EC 1.6.5.3)	1.510	1.324	1.181
cDNA FLJ53665, highly similar to Four and a half LIM domains protein 1;Four and a half LIM domains 1;Four and a half LIM domains protein 1;Skeletal muscle LIM-protein 1	1.500	1.168	1.031
Putative adenosylhomocysteinase 3;S-adenosylhomocysteine hydrolase-like protein 2;S-adenosyl-L-homocysteine hydrolase 3	1.499	1.165	1.031
28S ribosomal protein S9, mitochondrial	1.489	1.198	1.121
150 kDa oxygen-regulated protein;170 kDa glucose-regulated protein;Hypoxia up-regulated protein 1;cDNA FLJ54708, highly	1.480	1.179	1.041

similar to 150 kDa oxygen-regulated protein (Orp150)			
Complex I-39kD;NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 9, mitochondrial;NADH-ubiquinone oxidoreductase 39 kDa subunit	1.466	1.152	1.083
DDAHI;Dimethylargininase-1;N(G),N(G)-dimethylarginine dimethylaminohydrolase 1;cDNA FLJ54083, highly similar to NG,NG-dimethylarginine dimethylaminohydrolase 1 (EC 3.5.3.18);cDNA FLJ54119, highly similar to NG,NG-dimethylarginine dimethylaminohydrolase 1 (EC 3.5.3.18)	1.465	1.177	1.072
Phenylalanine--tRNA ligase beta chain;Phenylalanyl-tRNA synthetase beta chain	1.462	1.154	1.080
ABP-278;ABP-280 homolog;Actin-binding-like protein;Beta-filamin;Filamin homolog 1;Filamin-3;Filamin-B;Thyroid autoantigen;Truncated actin-binding protein	1.462	1.166	1.032
GTP-specific succinyl-CoA synthetase subunit beta;Succinyl-CoA ligase [GDP-forming] subunit beta, mitochondrial;Succinyl-CoA synthetase beta-G chain	1.459	1.151	1.018
TPPP/p20;Tubulin polymerization-promoting protein family member 3	1.455	1.376	1.275
F8W031;F8VXJ7;F8VP03	1.450	1.187	1.058
Protein NipSnap homolog 1	1.438	1.119	0.993
78 kDa gastrin-binding protein;Long chain 3-hydroxyacyl-CoA dehydrogenase;Long-chain enoyl-CoA hydratase;TP-alpha;Trifunctional enzyme subunit alpha, mitochondrial	1.434	1.118	1.032
Antioxidant enzyme AOE372;Peroxiredoxin IV;Peroxiredoxin-4;Thioredoxin peroxidase AO372;Thioredoxin-dependent peroxide reductase A0372	1.421	1.144	1.011
Calumenin;Crocalbin;IEF SSP 9302	1.418	1.286	1.136
GTPase IMAP family member 4;Immunity-associated nucleotide 1 protein;Immunity-associated protein 4;cDNA FLJ51351, highly similar to GTPase, IMAP family member 4	1.418	1.105	0.975
Plakophilin-2;Truncated plakophilin-2	1.417	1.103	1.000
Adaptor protein complex AP-1 mu-2 subunit;Adaptor-related protein complex 1 mu-2 subunit;AP-1 complex subunit mu-2;AP-mu chain family member mu1B;Clathrin assembly protein complex 1 medium chain 2;Golgi adaptor HA1/AP1 adaptin mu-2 subunit;Mu1B-adaptin;Mu-adaptin 2	1.417	1.262	1.117
Complex III subunit 1;Core protein I;Cytochrome b-c1 complex subunit 1, mitochondrial;Ubiquinol-cytochrome-c reductase complex core protein 1	1.413	1.209	1.071
90 kDa ribosomal protein S6 kinase 3;Insulin-stimulated protein kinase 1;MAP kinase-activated protein kinase 1b;pp90RSK2;Ribosomal protein S6 kinase alpha-3;Ribosomal S6 kinase 2;cDNA, FLJ79381, highly similar to Ribosomal protein S6 kinase alpha-3 (EC 2.7.11.1);cDNA FLJ56618, highly similar to Ribosomal protein S6 kinase alpha-3 (EC 2.7.11.1)	1.410	1.200	1.080
3-5 RNA exonuclease OLD35;PNPase old-35;Polynucleotide	1.407	1.255	1.155

phosphorylase 1;Polynucleotide phosphorylase-like protein;Polyribonucleotide nucleotidyltransferase 1, mitochondrial			
Complex I-B15;NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 4;NADH-ubiquinone oxidoreductase B15 subunit;Putative uncharacterized protein NDUFB4	1.397	1.113	1.010
Carnitine O-palmitoyltransferase 2, mitochondrial;Carnitine palmitoyltransferase II	1.396	1.089	0.976
Aldehyde dehydrogenase 5;Aldehyde dehydrogenase family 1 member B1;Aldehyde dehydrogenase X, mitochondrial;cDNA FLJ51238, highly similar to Aldehyde dehydrogenase X, mitochondrial (EC 1.2.1.3)	1.393	1.120	0.988
Coatomer subunit epsilon;Epsilon-coat protein;Coatomer protein complex, subunit epsilon, isoform CRA_e;Putative uncharacterized protein COPE	1.388	1.080	0.959
Ethylmalonic encephalopathy protein 1;Hepatoma subtracted clone one protein;Protein ETHE1, mitochondrial	1.383	1.090	0.974
SRA stem-loop-interacting RNA-binding protein, mitochondrial	1.382	1.075	0.960
15-hydroxyprostaglandin dehydrogenase [NADP+];Carbonyl reductase [NADPH] 1;NADPH-dependent carbonyl reductase 1;Prostaglandin 9-ketoreductase;Prostaglandin-E(2) 9-reductase;Putative uncharacterized protein CBR1;Carbonyl reductase 1, isoform CRA_c;cDNA FLJ60474, highly similar to Carbonyl reductase	1.379	1.085	1.043
ER-Golgi intermediate compartment 53 kDa protein;Gp58;Intracellular mannose-specific lectin MR60;Lectin mannose-binding 1;Protein ERGIC-53	1.374	1.129	1.046
Intestinal trefoil factor;Polypeptide P1.B;Trefoil factor 3	1.369	1.093	0.971
78 kDa glucose-regulated protein;Endoplasmic reticulum lumenal Ca(2+)-binding protein grp78;Heat shock 70 kDa protein 5;Immunoglobulin heavy chain-binding protein	1.366	1.104	1.046
Complex I-13kD-A;NADH dehydrogenase [ubiquinone] iron-sulfur protein 6, mitochondrial;NADH-ubiquinone oxidoreductase 13 kDa-A subunit	1.365	1.090	0.968
3-ketoacyl-CoA thiolase, mitochondrial;Acetyl-CoA acyltransferase;Beta-ketothiolase;Mitochondrial 3-oxoacyl-CoA thiolase;T1	1.365	1.097	1.028
Endoplasmic reticulum resident protein 46;Thioredoxin domain-containing protein 5;Thioredoxin-like protein p46;TXNDC5 protein;cDNA, FLJ96678, Homo sapiens thioredoxin domain containing 5 (TXNDC5), mRNA;HCG1811539, isoform CRA_b	1.363	1.103	0.994
Elongation factor Tu, mitochondrial;P43	1.361	1.065	0.985
Outer mitochondrial membrane protein porin 2;Voltage-dependent anion-selective channel protein 2;Voltage-dependent anion channel 2;cDNA FLJ60120, highly similar to Voltage-dependent anion-selective channel protein 2;cDNA, FLJ78818, highly similar to Voltage-dependent anion-selective channel	1.361	1.060	0.938

protein 2			
63 kDa membrane protein;Cytoskeleton-associated protein 4	1.361	1.102	0.975
Cytovillin;Ezrin;p81;Villin-2	1.360	1.057	0.937
Myosin I beta;Myosin-Ic	1.359	1.071	0.948
250/210 kDa paraneoplastic pemphigus antigen;Desmoplakin	1.356	1.126	1.002
Very long-chain specific acyl-CoA dehydrogenase, mitochondrial	1.355	1.089	0.983
15-oxoprostaglandin 13-reductase;Prostaglandin reductase 2;Zinc-binding alcohol dehydrogenase domain-containing protein 1	1.353	1.160	1.057
Complex III subunit 2;Core protein II;Cytochrome b-c1 complex subunit 2, mitochondrial;Ubiquinol-cytochrome-c reductase complex core protein 2	1.345	1.200	1.182
Aspartate aminotransferase, mitochondrial;Fatty acid-binding protein;Glutamate oxaloacetate transaminase 2;Plasma membrane-associated fatty acid-binding protein;Transaminase A	1.339	1.041	0.919
CML33;Phenylalanine--tRNA ligase alpha chain;Phenylalanyl-tRNA synthetase alpha chain;cDNA FLJ50378, highly similar to Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20)	1.337	1.179	1.051
Sodium pump subunit alpha-1;Sodium/potassium-transporting ATPase subunit alpha-1;ATPase, Na+/K+ transporting, alpha 1 polypeptide, isoform CRA_a;cDNA FLJ52430, highly similar to Sodium/potassium-transporting ATPase alpha-1 chain (EC 3.6.3.9)	1.335	1.042	0.927
Putative uncharacterized protein MLLT4;Afadin;ALL1-fused gene from chromosome 6 protein;Myeloid/lymphoid or mixed-lineage leukemia (Trithorax homolog, Drosophila); translocated to, 4	1.334	1.277	1.207
Cytochrome c oxidase polypeptide Vb;Cytochrome c oxidase subunit 5B, mitochondrial	1.332	1.229	1.092
35 kDa lectin;Carbohydrate-binding protein 35;Galactose-specific lectin 3;Galactoside-binding protein;Galectin-3;IgE-binding protein;L-31;Laminin-binding protein;Lectin L-29;Mac-2 antigen	1.328	1.057	0.972
Complex I-B22;LYR motif-containing protein 3;NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 9;NADH-ubiquinone oxidoreductase B22 subunit	1.327	1.179	1.045
3-ketoacyl-CoA thiolase;Acetyl-CoA acyltransferase;Beta-ketothiolase;TP-beta;Trifunctional enzyme subunit beta, mitochondrial;cDNA FLJ56214, highly similar to Trifunctional enzyme subunit beta, mitochondrial;Putative uncharacterized protein HADHB	1.325	1.030	0.972
Endoplasmic reticulum resident protein 28;Endoplasmic reticulum resident protein 29;Endoplasmic reticulum resident protein 31	1.322	1.061	0.938
Alu corepressor 1;Antioxidant enzyme B166;Liver tissue 2D-page spot 71B;Peroxiredoxin V;Peroxiredoxin-5, mitochondrial;Peroxisomal antioxidant enzyme;PLP;Thioredoxin peroxidase PMP20;Thioredoxin reductase;TPx type VI;Putative uncharacterized protein PRDX5	1.316	1.071	0.947

ER-Golgi SNARE of 24 kDa;SEC22 vesicle-trafficking protein homolog B;SEC22 vesicle-trafficking protein-like 1;Vesicle-trafficking protein SEC22b	1.315	1.023	0.937
Calcium-binding mitochondrial carrier protein Aralar2;Citrin;Mitochondrial aspartate glutamate carrier 2;Solute carrier family 25 member 13	1.314	1.021	0.944
RRP12-like protein	1.311	1.083	0.961
Endoplasmic reticulum resident protein 70;Endoplasmic reticulum resident protein 72;Protein disulfide-isomerase A4	1.311	1.110	0.980
Myosin-Id	1.308	1.070	0.944
Actin-depolymerizing factor;Destrin	1.306	1.027	0.918
Complex I-B14.5a;NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 7;NADH-ubiquinone oxidoreductase subunit B14.5a	1.305	1.115	1.169
Beta-G1;Beta-glucuronidase	1.297	1.019	0.921
Chymotrypsin-like elastase family member 3A;Elastase IIIA;Elastase-3A;Protease E	1.297	1.020	0.910
17-beta-hydroxysteroid dehydrogenase 11;17-beta-hydroxysteroid dehydrogenase XI;Cutaneous T-cell lymphoma-associated antigen HD-CL-03;Dehydrogenase/reductase SDR family member 8;Estradiol 17-beta-dehydrogenase 11;Retinal short-chain dehydrogenase/reductase 2	1.294	1.013	0.895
Interleukin-25;Stromal cell-derived growth factor SF20;UPF0556 protein C19orf10	1.292	1.121	0.990
Complex III subunit 3;Complex III subunit III;Cytochrome b;Cytochrome b-c1 complex subunit 3;Ubiquinol-cytochrome-c reductase complex cytochrome b subunit	1.287	1.237	1.209
Cellular thyroid hormone-binding protein;p55;Prolyl 4-hydroxylase subunit beta;Protein disulfide-isomerase;cDNA FLJ59430, highly similar to Protein disulfide-isomerase (EC 5.3.4.1)	1.285	1.018	0.995
Importin-4;Importin-4b;Ran-binding protein 4	1.280	1.268	1.137
94 kDa glucose-regulated protein;Endoplasmin;gp96 homolog;Heat shock protein 90 kDa beta member 1;Tumor rejection antigen 1	1.277	1.040	1.005
Amine oxidase [flavin-containing] A;Monoamine oxidase type A;cDNA FLJ61220, highly similar to Amine oxidase (flavin-containing) A (EC 1.4.3.4)	1.276	1.023	0.903
Ubiquitin-fold modifier 1	1.269	1.042	1.089
Antigen NY-CO-4;Elongation factor 1-delta	1.267	1.160	1.025
3-phosphoadenosine-5-phosphosulfate synthase;Adenosine-5-phosphosulfate 3-phosphotransferase;Adenylylsulfate 3-phosphotransferase;Adenylyl-sulfate kinase;APS kinase;ATP-sulfurylase;Bifunctional 3-phosphoadenosine 5-phosphosulfate synthase 2;Sulfate adenylate transferase;Sulfate adenylyltransferase;Sulfurylase kinase 2	1.263	0.990	0.874
Alpha-adducin;Erythrocyte adducin subunit alpha;Adducin 1	1.259	1.110	0.981

(Alpha);Adducin 1 (Alpha), isoform CRA_e;ADD1 protein			
Microsomal signal peptidase 25 kDa subunit;Signal peptidase complex subunit 2	1.257	0.978	0.863
Quiescin Q6;Sulfhydryl oxidase 1	1.257	1.111	1.053
Acetoacetyl-CoA thiolase;Acetyl-CoA acetyltransferase, mitochondrial;T2	1.255	0.994	0.887
2-oxoglutarate dehydrogenase complex component E1;2-oxoglutarate dehydrogenase, mitochondrial;Alpha-ketoglutarate dehydrogenase	1.254	0.996	0.933
Complex III subunit 7;Complex III subunit VII;Cytochrome b-c1 complex subunit 7;QP-C;Ubiquinol-cytochrome c reductase complex 14 kDa protein;cDNA FLJ52271, moderately similar to Ubiquinol-cytochrome c reductase complex 14 kDa protein (EC 1.10.2.2)	1.251	1.078	0.966
Calcium-activated chloride channel family member 1;Calcium-activated chloride channel protein 1;Calcium-activated chloride channel regulator 1	1.250	0.983	0.905
Complement component 4A (Rodgers blood group);Putative uncharacterized protein C4A;Complement component C4B (Childo blood group);Complement component C4B (Childo blood group) 2;C4B1;Complement component 4B (Childo blood group)	1.249	0.982	0.981
Actin-interacting protein 1;NORI-1;WD repeat-containing protein 1;cDNA FLJ58303, highly similar to WD repeat protein 1	1.249	1.050	1.050
Catalase	1.245	0.986	0.951
Proteasome subunit alpha type-7;Proteasome subunit RC6-1;Proteasome subunit XAPC7;Proteasome subunit alpha type	1.244	0.988	1.122
Heat shock-related 70 kDa protein 2;cDNA FLJ40505 fis, clone TESTI2045562, highly similar to HEAT SHOCK-RELATED 70 kDa PROTEIN 2	1.244	1.032	0.922
Endopeptidase SP18;Microsomal signal peptidase 18 kDa subunit;SEC11 homolog A;SEC11-like protein 1;Signal peptidase complex catalytic subunit SEC11A;SPC18;cDNA FLJ51313, highly similar to Microsomal signal peptidase 18 kDa subunit(EC 3.4.-.-);SEC11-like 1 (S. cerevisiae), isoform CRA_d	1.240	1.009	1.057
Inorganic pyrophosphatase;Pyrophosphate phospho-hydrolase;Pyrophosphatase (Inorganic) 1	1.238	0.962	0.998
Myosin heavy chain 11;Myosin heavy chain, smooth muscle isoform;Myosin-11;SMMHC;Myosin heavy chain 11 smooth muscle isoform	1.234	0.963	0.862
Clathrin heavy chain 1;Clathrin heavy chain on chromosome 17	1.232	0.982	0.869
Villin-1;cDNA FLJ57609, highly similar to Villin-1	1.228	0.979	0.871
Centromere protein V;Nuclear protein p30;Proline-rich protein 6	1.227	1.172	1.067
Cathepsin C;Cathepsin J;Dipeptidyl peptidase 1;Dipeptidyl peptidase 1 exclusion domain chain;Dipeptidyl peptidase 1 heavy chain;Dipeptidyl peptidase 1 light chain;Dipeptidyl peptidase I;Dipeptidyl peptidase I exclusion domain chain;Dipeptidyl peptidase I heavy chain;Dipeptidyl peptidase I light	1.224	0.962	0.855

chain;Dipeptidyl transferase			
Hydroxyacyl-coenzyme A dehydrogenase, mitochondrial;Medium and short-chain L-3-hydroxyacyl-coenzyme A dehydrogenase;Short-chain 3-hydroxyacyl-CoA dehydrogenase	1.224	1.009	0.903
Aldo-keto reductase family 1 member B10;Aldose reductase-like;Aldose reductase-related protein;ARL-1;Small intestine reductase	1.223	0.954	0.851
Carbonate dehydratase II;Carbonic anhydrase 2;Carbonic anhydrase C;Carbonic anhydrase II	1.223	1.031	0.910
Putative uncharacterized protein PRRC1;Protein PRRC1	1.223	0.951	0.938
Cytosolic malate dehydrogenase;Malate dehydrogenase, cytoplasmic;Malate dehydrogenase;Putative uncharacterized protein MDH1	1.222	0.988	0.875
Cadherin-associated Src substrate;Catenin delta-1;p120 catenin;p120(cas);Putative uncharacterized protein CTNND1	1.222	1.013	0.894
Metavinculin;Vinculin	1.221	0.977	0.955
Amphiregulin-associated protein;Midgestation and kidney protein;Midkine;Neurite outgrowth-promoting factor 2;Neurite outgrowth-promoting protein	1.220	1.206	1.070
Putative uncharacterized protein APEH;Acylamino-acid-releasing enzyme;Acylaminoacyl-peptidase;Acyl-peptide hydrolase;Oxidized protein hydrolase	1.219	0.973	0.871
Kallikrein inhibitor;Kallistatin;Peptidase inhibitor 4;Serpin A4	1.218	1.258	1.166
Dihydrolipoamide dehydrogenase;Dihydrolipoyl dehydrogenase, mitochondrial;Glycine cleavage system L protein;cDNA FLJ50515, highly similar to Dihydrolipoyl dehydrogenase, mitochondrial (EC 1.8.1.4);Dihydrolipoyl dehydrogenase	1.213	0.980	0.911
5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase;AICAR transformylase;ATIC;Bifunctional purine biosynthesis protein PURH;IMP cyclohydrolase;IMP synthase;Inosinicase;Phosphoribosylaminoimidazolecarboxamide formyltransferase	1.212	0.943	0.864
Glioma pathogenesis-related protein 2;Golgi-associated plant pathogenesis-related protein 1;GLI pathogenesis-related 2	1.210	1.027	0.947
Signal transducer and activator of transcription 1-alpha/beta;Transcription factor ISGF-3 components p91/p84	1.209	1.167	1.046
PDHE1-A type I;Pyruvate dehydrogenase E1 component subunit alpha, somatic form, mitochondrial	1.205	0.948	0.932
Glucose phosphomutase 1;Phosphoglucomutase-1;cDNA FLJ50606, highly similar to Phosphoglucomutase-1 (EC 5.4.2.2)	1.202	0.935	0.998
18 kDa Alu RNA-binding protein;Signal recognition particle 14 kDa protein	1.200	1.015	0.912
Isovaleryl-CoA dehydrogenase, mitochondrial;cDNA FLJ16602 fis, clone TESTI4007816, highly similar to Isovaleryl-CoA dehydrogenase, mitochondrial (EC 1.3.99.10);Isovaleryl Coenzyme A dehydrogenase, isoform CRA_b	1.198	1.177	1.045
Cullin-associated and neddylation-dissociated protein 1;Cullin-	1.197	1.039	0.994

associated NEDD8-dissociated protein 1;p120 CAND1;TBP-interacting protein of 120 kDa A			
Cytochrome c oxidase polypeptide VIc;Cytochrome c oxidase subunit 6C	1.196	1.056	0.935
Beta-hexosaminidase subunit alpha;Beta-N-acetylhexosaminidase subunit alpha;N-acetyl-beta-glucosaminidase subunit alpha	1.194	1.105	1.110
HCNPpp;Hippocampal cholinergic neurostimulating peptide;Neuropolypeptide h3;Phosphatidylethanolamine-binding protein 1;Prostatic-binding protein;Raf kinase inhibitor protein;cDNA FLJ51535, highly similar to Phosphatidylethanolamine-binding protein 1	1.193	0.958	0.996
59 kDa serine/threonine-protein kinase;ILK-1;ILK-2;Integrin-linked protein kinase;p59ILK;cDNA FLJ50979, moderately similar to Integrin-linked protein kinase (EC 2.7.11.1);cDNA FLJ53825, highly similar to Integrin-linked protein kinase 1 (EC 2.7.11.1)	1.192	1.090	0.965
Collagen alpha-1(XV) chain;Endostatin;Endostatin-XV;Related to endostatin;Restin	1.188	0.994	0.894
Desmoyokin;Neuroblast differentiation-associated protein AHNAK	1.187	0.968	0.893
HBeAg-binding protein 2 binding protein A;Mannose-P-dolichol utilization defect 1 protein;Suppressor of Lec15 and Lec35 glycosylation mutation homolog;My008 protein;cDNA FLJ57793, moderately similar to Mannose-P-dolichol utilization defect 1 protein;cDNA FLJ14836 fis, clone OVARC1001702	1.183	0.933	0.904
Outer mitochondrial membrane protein porin 1;Plasmalemmal porin;Porin 31HL;Porin 31HM;Voltage-dependent anion-selective channel protein 1	1.182	0.920	0.858
Alpha E-catenin;Cadherin-associated protein;Catenin alpha-1;Renal carcinoma antigen NY-REN-13;cDNA FLJ54047, highly similar to Alpha-1 catenin (Cadherin-associated protein);Catenin (Cadherin-associated protein), alpha 1, 102kDa, isoform CRA_c;CTNNA1 protein	1.180	0.933	0.847
Antigen KI-67	1.180	1.623	1.483
Glutamate dehydrogenase 1, mitochondrial;cDNA FLJ55203, highly similar to Glutamate dehydrogenase 1, mitochondrial (EC 1.4.1.3);cDNA FLJ16138 fis, clone BRALZ2017531, highly similar to Glutamate dehydrogenase 1, mitochondrial (EC 1.4.1.3);Glutamate dehydrogenase 1, isoform CRA_a;Glutamate dehydrogenase 2, mitochondrial	1.179	0.981	0.896
Complex I-PDSW;NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 10;NADH-ubiquinone oxidoreductase PDSW subunit;NADH dehydrogenase (Ubiquinone) 1 beta subcomplex, 10, 22kDa, isoform CRA_a;NDUFB10 protein	1.176	1.263	1.126
Hydroxysteroid dehydrogenase-like protein 2;cDNA FLJ61200, highly similar to Homo sapiens hydroxysteroid dehydrogenase like 2 (HSDL2), mRNA	1.175	0.941	0.868

Elongation factor Ts, mitochondrial;Elongation factor Ts	1.169	1.294	1.165
5H9 antigen;CD9 antigen;Cell growth-inhibiting gene 2 protein;Leukocyte antigen MIC3;Motility-related protein;p24;Tetraspanin-29;Putative uncharacterized protein CD9;cDNA FLJ51032, highly similar to CD9 antigen	1.169	0.916	0.826
180 kDa ribosome receptor homolog;ES/130-related protein;Ribosome receptor protein;Ribosome-binding protein 1	1.168	0.928	1.046
Hematopoietic cell-specific LYN substrate 1;Hematopoietic lineage cell-specific protein;LckBP1;p75	1.167	1.083	1.017
DDAHII;Dimethylargininase-2;N(G),N(G)-dimethylarginine dimethylaminohydrolase 2;Protein G6a;S-phase protein;Dimethylarginine dimethylaminohydrolase 2	1.166	0.968	0.901
Glutathione S-transferase omega-1;Glutathione S-transferase omega 1, isoform CRA_a;Glutathione S-transferase omega 1	1.162	0.906	0.890
Apoptotic chromatin condensation inducer in the nucleus	1.160	1.040	0.920
Aldehyde dehydrogenase family 6 member A1;Methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial	1.158	0.938	0.845
Protein disulfide isomerase P5;Protein disulfide-isomerase A6;Thioredoxin domain-containing protein 7;cDNA FLJ58502, highly similar to Protein disulfide-isomerase A6 (EC 5.3.4.1)	1.158	0.924	0.953
Nuclear transport factor 2;Placental protein 15	1.157	1.016	0.914
Complex I-23kD;NADH dehydrogenase [ubiquinone] iron-sulfur protein 8, mitochondrial;NADH-ubiquinone oxidoreductase 23 kDa subunit;TYKY subunit	1.151	1.148	1.025
Complex III subunit 5;Complex III subunit IX;Cytochrome b-c1 complex subunit 11;Cytochrome b-c1 complex subunit 5;Cytochrome b-c1 complex subunit Rieske, mitochondrial;Rieske iron-sulfur protein;Ubiquinol-cytochrome c reductase 8 kDa protein;Ubiquinol-cytochrome c reductase iron-sulfur subunit;Putative cytochrome b-c1 complex subunit Rieske-like protein 1	1.150	1.059	0.936
Nicotinamide phosphoribosyltransferase;Pre-B-cell colony-enhancing factor 1;Visfatin;Novel protein similar to Pre-B cell enhancing factor (PBEF)	1.148	0.893	0.823
Basic leucine zipper and W2 domain-containing protein 2;Putative uncharacterized protein BZW2;Basic leucine zipper and W2 domains 2, isoform CRA_b	1.147	0.987	0.974
ATP-specific succinyl-CoA synthetase subunit beta;Renal carcinoma antigen NY-REN-39;Succinyl-CoA ligase [ADP-forming] subunit beta, mitochondrial;Succinyl-CoA synthetase beta-A chain;Succinate-CoA ligase, ADP-forming, beta subunit	1.147	0.928	0.922
3-hydroxyisobutyryl-CoA hydrolase, mitochondrial;3-hydroxyisobutyryl-coenzyme A hydrolase	1.145	1.152	1.058
E3 UFM1-protein ligase 1	1.143	1.066	1.053
UPF0197 transmembrane protein C11orf10	1.140	0.886	0.782
Interferon-induced protein 53;T1-TrpRS;T2-TrpRS;Tryptophan--tRNA ligase;Tryptophanyl-tRNA synthetase, cytoplasmic	1.139	1.201	1.084

Tumor protein D52-like 2;Tumor protein D52-like 2, isoform CRA_e;Tumor protein D54	1.136	1.002	0.885
Glycogen phosphorylase, brain form	1.131	0.884	0.787
Cytosolic thyroid hormone-binding protein;Opa-interacting protein 3;p58;Pyruvate kinase 2/3;Pyruvate kinase isozymes M1/M2;Pyruvate kinase muscle isozyme;Thyroid hormone-binding protein 1;Tumor M2-PK	1.131	1.043	0.931
58 kDa glucose-regulated protein;58 kDa microsomal protein;Disulfide isomerase ER-60;Endoplasmic reticulum resident protein 57;Endoplasmic reticulum resident protein 60;Protein disulfide-isomerase A3;cDNA PSEC0175 fis, clone OVARC1000169, highly similar to Protein disulfide-isomerase A3 (EC 5.3.4.1)	1.129	0.910	1.022
Glutathione S-transferase kappa 1;Glutathione S-transferase subunit 13;GST 13-13;GST class-kappa;GSTK1-1	1.128	0.978	0.995
Enoyl-CoA hydratase 1;Enoyl-CoA hydratase, mitochondrial;Short-chain enoyl-CoA hydratase	1.125	0.890	0.787
Active breakpoint cluster region-related protein;cDNA FLJ54747, highly similar to Active breakpoint cluster region-related protein	1.124	0.918	0.913
HLA-DR-associated protein II;Inhibitor of granzyme A-activated DNase;PHAPII;Phosphatase 2A inhibitor I2PP2A;Protein SET;Template-activating factor I;SET nuclear oncogene;Putative uncharacterized protein SET	1.118	1.145	1.099
Tubulin beta-2 chain;Tubulin beta-2C chain	1.117	0.900	0.805
Beta-II spectrin;Fodrin beta chain;Spectrin beta chain, brain 1;Spectrin, non-erythroid beta chain 1	1.117	0.905	0.800
Cyclophilin B;CYP-S1;Peptidyl-prolyl cis-trans isomerase B;Rotamase B;S-cyclophilin	1.115	0.920	0.903
ATP synthase subunit a;F-ATPase protein 6	1.112	0.888	0.826
Putative uncharacterized protein ARPC4;Actin-related protein 2/3 complex subunit 4;Arp2/3 complex 20 kDa subunit	1.106	0.931	0.919
Carboxymethylenebutenolidase homolog	1.106	0.927	0.820
Vasodilator-stimulated phosphoprotein	1.106	1.087	0.995
47 kDa mannose 6-phosphate receptor-binding protein;Cargo selection protein TIP47;Mannose-6-phosphate receptor-binding protein 1;Perilipin-3;Placental protein 17	1.102	0.978	0.917
All-trans-13,14-dihydroretinol saturase;All-trans-retinol 13,14-reductase	1.099	1.249	1.121
Beta-coat protein;Coatomer subunit beta;p102;cDNA FLJ56271, highly similar to Coatomer subunit beta;Coatomer protein complex, subunit beta 2 (Beta prime), isoform CRA_b	1.097	0.880	0.905
BPG-dependent PGAM 1;Phosphoglycerate mutase 1;Phosphoglycerate mutase isozyme B	1.096	1.042	0.940
Cadherin family member 5;Desmoglein-2;HDGC	1.096	1.083	1.010
Superoxide dismutase [Mn], mitochondrial;Superoxide dismutase	1.095	0.871	0.872
Interferon-induced 15 kDa protein;Interferon-induced 17 kDa	1.093	0.918	0.810

protein;Ubiquitin cross-reactive protein			
Transmembrane and coiled-coil domain-containing protein 1;Transmembrane and coiled-coil domains protein 4;Xenogeneic cross-immune protein PCIA3;Putative uncharacterized protein TMCO1	1.092	0.849	0.786
Actin-binding protein 280;Alpha-filamin;Endothelial actin-binding protein;Filamin-1;Filamin-A;Non-muscle filamin;Filamin A, alpha (Actin binding protein 280)	1.091	0.939	0.851
ABP-280-like protein;ABP-L;Actin-binding-like protein;Filamin-2;Filamin-C;Gamma-filamin	1.089	0.855	0.761
1-acylglycerophosphocholine O-acyltransferase;1-acylglycerophosphoserine O-acyltransferase;Lysophosphatidylcholine acyltransferase;Lysophosphatidylcholine acyltransferase 3;Lysophosphatidylserine acyltransferase;Lysophospholipid acyltransferase 5;Membrane-bound O-acyltransferase domain-containing protein 5;cDNA FLJ55747, highly similar to Membrane bound O-acyltransferase domain-containing protein 5 (EC 2.3.-.-)	1.088	1.242	1.118
Alcohol dehydrogenase 1C;Alcohol dehydrogenase subunit gamma	1.086	0.976	0.875
Beta-hexosaminidase subunit beta;Beta-hexosaminidase subunit beta chain A;Beta-hexosaminidase subunit beta chain B;Beta-N-acetylhexosaminidase subunit beta;Cervical cancer proto-oncogene 7 protein;N-acetyl-beta-glucosaminidase subunit beta;ENC-1AS	1.081	0.841	0.811
E3 ubiquitin/ISG15 ligase TRIM25;Estrogen-responsive finger protein;RING finger protein 147;Tripartite motif-containing protein 25;Ubiquitin/ISG15-conjugating enzyme TRIM25;Zinc finger protein 147	1.079	0.982	0.925
p195;Ras GTPase-activating-like protein IQGAP1	1.078	0.877	0.859
Cytochrome c oxidase polypeptide IV;Cytochrome c oxidase subunit 4 isoform 1, mitochondrial;Cytochrome c oxidase subunit IV isoform 1;COX4I1 protein	1.074	1.097	0.969
Intramembrane protease 1;Minor histocompatibility antigen H13;Presenilin-like protein 3;Signal peptide peptidase	1.074	0.847	0.768
Complex I-ASHI;NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 8, mitochondrial;NADH-ubiquinone oxidoreductase ASHI subunit;NADH dehydrogenase (Ubiquinone) 1 beta subcomplex, 8, 19kDa;NADH dehydrogenase (Ubiquinone) 1 beta subcomplex, 8, 19kDa, isoform CRA_a;cDNA FLJ52503, highly similar to NADH dehydrogenase (ubiquinone) 1 beta subcomplex subunit 8, mitochondrial (EC 1.6.5.3) (EC 1.6.99.3) (NADH-ubiquinone oxidoreductase ASHI subunit) (Complex I-ASHI) (CI-ASHI)	1.073	0.986	1.194
Cathepsin D;Cathepsin D heavy chain;Cathepsin D light chain	1.070	0.864	0.790
Golgi transport 1 homolog B;hGOT1a;Putative NF-kappa-B-activating protein 470;Vesicle transport protein GOT1B	1.069	0.834	0.763

ADP-ribosylation factor 4;Putative uncharacterized protein ARF4	1.069	0.890	0.832
Calnexin;IP90;Major histocompatibility complex class I antigen-binding protein p88;p90;cDNA FLJ54242, highly similar to Calnexin	1.068	0.857	0.983
Macropain subunit C8;Multicatalytic endopeptidase complex subunit C8;Proteasome component C8;Proteasome subunit alpha type-3	1.066	0.845	0.953
Endoplasmic oxidoreductin-1-like protein;ERO1-like protein alpha;Oxidoreductin-1-L-alpha	1.066	0.875	0.912
Elastin microfibril interface-located protein 1;EMILIN-1	1.064	0.846	0.788
Membrane protein p24A;Transmembrane emp24 domain-containing protein 2;cDNA FLJ52153, highly similar to Transmembrane emp24 domain-containing protein 2	1.059	0.828	0.826
60 kDa SS-A/Ro ribonucleoprotein;Ro 60 kDa autoantigen;Sjoegren syndrome antigen A2;Sjoegren syndrome type A antigen;TROVE domain family member 2;TROVE domain family, member 2;TROVE domain family, member 2, isoform CRA_c;TROVE domain family, member 2, isoform CRA_e;TROVE domain family, member 2, isoform CRA_d	1.057	1.024	0.928
Serine/threonine-protein phosphatase PP1-beta catalytic subunit	1.049	0.850	0.916
Complex I-49kD;NADH dehydrogenase [ubiquinone] iron-sulfur protein 2, mitochondrial;NADH-ubiquinone oxidoreductase 49 kDa subunit;cDNA, FLJ78876, highly similar to NADH-ubiquinone oxidoreductase 49 kDa subunit, mitochondrial (EC 1.6.5.3)	1.048	0.883	0.931
Glycoprotein GP36b;Lectin mannose-binding 2;Vesicular integral-membrane protein VIP36;cDNA FLJ52285, highly similar to Vesicular integral-membrane protein VIP36	1.039	0.809	0.779
Azoreductase;DT-diaphorase;Menadione reductase;NAD(P)H dehydrogenase [quinone] 1;NAD(P)H:quinone oxidoreductase 1;Phylloquinone reductase;Quinone reductase 1;cDNA FLJ50573, highly similar to Homo sapiens NAD(P)H dehydrogenase, quinone 1 (NQO1), transcript variant 3, mRNA	1.039	0.880	0.848
Fortilin;Histamine-releasing factor;p23;Translationally-controlled tumor protein;TPT1 protein;Tumor protein, translationally-controlled 1;Tumor protein, translationally-controlled 1, isoform CRA_a	1.037	0.854	0.770
Adapter protein CMS;Cas ligand with multiple SH3 domains;CD2-associated protein	1.033	1.063	0.962
Oxysterol-binding protein 1	1.032	1.125	1.001
B5;Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3A;Integral membrane protein 1;Transmembrane protein TMC	1.029	0.966	0.888
11S regulator complex subunit alpha;Activator of multicatalytic protease subunit 1;Interferon gamma up-regulated I-5111 protein;Proteasome activator 28 subunit alpha;Proteasome activator complex subunit 1;Putative uncharacterized protein PSME1	1.028	1.069	0.953

CFR-1;Cysteine-rich fibroblast growth factor receptor;E-selectin ligand 1;Golgi apparatus protein 1;Golgi sialoglycoprotein MG-160	1.028	0.879	0.890
Ubiquitin carrier protein D3;Ubiquitin-conjugating enzyme E2 D3;Ubiquitin-conjugating enzyme E2(17)KB 3;Ubiquitin-conjugating enzyme E2-17 kDa 3;Ubiquitin-protein ligase D3;Ubiquitin carrier protein D2;Ubiquitin-conjugating enzyme E2 D2;Ubiquitin-conjugating enzyme E2(17)KB 2;Ubiquitin-conjugating enzyme E2-17 kDa 2;Ubiquitin-protein ligase D2;Ubiquitin carrier protein	1.028	0.940	0.831
Activated RNA polymerase II transcriptional coactivator p15;p14;Positive cofactor 4;SUB1 homolog	1.024	0.913	0.818
HLA-B-associated transcript 3;Large proline-rich protein BAT3;Protein G3;HLA-B associated transcript 3;HLA-B associated transcript 3, isoform CRA_a	1.024	0.942	0.836
Cytochrome c oxidase polypeptide II;Cytochrome c oxidase subunit 2	1.024	1.008	1.048
Histone H1;Histone H1(0);Histone H1.0	1.022	0.846	0.748
Echinoderm microtubule-associated protein-like 4;Restrictedly overexpressed proliferation-associated protein;Ropp 120;Putative uncharacterized protein EML4	1.021	0.899	0.822
Fructose-bisphosphate aldolase A;Lung cancer antigen NY-LU-1;Muscle-type aldolase	1.019	0.828	0.750
High density lipoprotein-binding protein;Vigilin	1.018	0.791	0.705
32 kDa accessory protein;Vacuolar proton pump subunit d 1;V-ATPase 40 kDa accessory protein;V-ATPase AC39 subunit;V-type proton ATPase subunit d 1	1.015	0.810	0.900
NADH dehydrogenase [ubiquinone] flavoprotein 2, mitochondrial;NADH-ubiquinone oxidoreductase 24 kDa subunit	1.015	0.833	0.956
Maternal-embryonic 3;Vacuolar protein sorting-associated protein 35;Vesicle protein sorting 35	1.013	0.795	0.951
Histone H3	1.013	0.825	0.729
17-beta-hydroxysteroid dehydrogenase type 2;20 alpha-hydroxysteroid dehydrogenase;E2DH;Estradiol 17-beta-dehydrogenase 2;Microsomal 17-beta-hydroxysteroid dehydrogenase;Testosterone 17-beta-dehydrogenase	1.013	1.457	1.303
56 kDa selenium-binding protein;Selenium-binding protein 1;cDNA FLJ61035, highly similar to Selenium-binding protein 1;Selenium binding protein 1	1.013	0.789	0.706
Cell proliferation-inducing gene 19 protein;LDH muscle subunit;L-lactate dehydrogenase A chain;Renal carcinoma antigen NY-REN-59	1.006	0.936	1.001
Tight junction protein 1;Tight junction protein ZO-1;Zona occludens protein 1;Zonula occludens protein 1	1.002	1.323	1.170
Ras-related protein Rab-18;RAB18, member RAS oncogene family	1.000	0.849	1.001
Epithelial protein lost in neoplasm;LIM domain and actin-binding protein 1;cDNA FLJ55990, highly similar to LIM domain and actin-	1.000	0.980	0.875

binding protein 1			
Iron regulatory protein 2;Iron-responsive element-binding protein 2	0.999	0.786	0.761
G protein subunit beta-2;Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-2;Transducin beta chain 2;Putative uncharacterized protein GNB2	0.995	0.799	0.785
40S ribosomal protein S13	0.991	0.857	0.891
Heat shock 70 kDa protein 1/2;Heat shock 70 kDa protein 1A/1B;Heat shock 70kDa protein 1A	0.989	1.199	1.082
LIM and SH3 domain protein 1;Metastatic lymph node gene 50 protein;Putative uncharacterized protein LASP1;cDNA FLJ51834, highly similar to LIM and SH3 domain protein 1;cDNA FLJ52195, highly similar to LIM and SH3 domain protein 1	0.981	0.977	0.872
60S ribosomal protein L26	0.981	0.765	0.972
20 kDa myosin light chain;MLC-2C;Myosin regulatory light chain 2, smooth muscle isoform;Myosin regulatory light chain 9;Myosin regulatory light chain MRLC1;Myosin regulatory light polypeptide 9;Myosin RLC	0.981	1.027	0.924
Guanine nucleotide-binding protein G(y) subunit alpha;Guanine nucleotide-binding protein subunit alpha-11	0.978	0.764	0.780
Meg-3;Niban-like protein 1;Protein FAM129B	0.972	0.829	0.789
ADP-ribosylation factor-like protein 6-interacting protein 5;Cytoskeleton-related vitamin A-responsive protein;Dermal papilla-derived protein 11;Glutamate transporter EAAC1-interacting protein;GTRAP3-18;JM5;PRA1 family protein 3;Prenylated Rab acceptor protein 2;Protein JWa;Putative MAPK-activating protein PM27	0.969	1.128	1.069
60S ribosomal protein L27	0.968	0.799	0.741
Beta-globin;Hemoglobin beta chain;Hemoglobin subunit beta;LVV-hemorphin-7	0.967	0.996	1.014
GDP-4-keto-6-deoxy-D-mannose-3,5-epimerase-4-reductase;GDP-L-fucose synthase;Protein FX;Red cell NADP(H)-binding protein;Short-chain dehydrogenase/reductase family 4E member 1	0.964	0.927	0.848
26S protease regulatory subunit 6A;26S proteasome AAA-ATPase subunit RPT5;Proteasome 26S subunit ATPase 3;Proteasome subunit P50;Tat-binding protein 1	0.963	1.125	1.059
Apolipoprotein A-I-binding protein;YjeF N-terminal domain-containing protein 1	0.958	1.024	0.970
APEX nuclease;Apurinic-apyrimidinic endonuclease 1;DNA-(apurinic or apyrimidinic site) lyase;Protein REF-1	0.954	1.016	0.900
Signal recognition particle 54 kDa protein	0.953	0.743	0.667
5F7;Basigin;Collagenase stimulatory factor;Extracellular matrix metalloproteinase inducer;Leukocyte activation antigen M6;OK blood group antigen;Tumor cell-derived collagenase stimulatory factor	0.952	0.794	0.849
AIR carboxylase;Multifunctional protein	0.952	0.740	0.673

ADE2;Phosphoribosylaminoimidazole carboxylase;Phosphoribosylaminoimidazole-succinocarboxamide synthase;SAICAR synthetase			
70 kDa peroxisomal membrane protein;ATP-binding cassette sub-family D member 3	0.951	0.926	0.853
Actin, aortic smooth muscle;Alpha-actin-2;Cell growth-inhibiting gene 46 protein;Actin, alpha 1, skeletal muscle	0.951	0.935	0.971
3-hydroxybutyrate dehydrogenase type 2;Dehydrogenase/reductase SDR family member 6;Oxidoreductase UCPA;R-beta-hydroxybutyrate dehydrogenase	0.946	0.753	0.896
Kinesin light chain 4;Kinesin-like protein 8;cDNA FLJ58264, highly similar to Kinesin light chain 4	0.945	1.267	1.218
Deubiquitinating enzyme 15;Ubiquitin carboxyl-terminal hydrolase 15;Ubiquitin thioesterase 15;Ubiquitin-specific-processing protease 15;Unph-2;Unph4	0.944	1.145	1.049
Putative uncharacterized protein KIAA0664;Protein KIAA0664	0.933	0.738	0.830
Protein SCO2 homolog, mitochondrial	0.928	0.744	0.657
35-alpha calcimedin;Annexin A3;Annexin III;Annexin-3;Inositol 1,2-cyclic phosphate 2-phosphohydrolase;Lipocortin III;Placental anticoagulant protein III	0.926	1.006	0.966
Inosine phosphorylase;Purine nucleoside phosphorylase	0.919	0.803	0.816
Complex I-B8;NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 2;NADH-ubiquinone oxidoreductase B8 subunit	0.916	0.719	0.780
Lamin-B2	0.914	0.712	0.910
Signal transducing adapter molecule 1;Putative uncharacterized protein STAM	0.906	0.965	0.883
Plasma membrane calcium ATPase isoform 1;Plasma membrane calcium pump isoform 1;Plasma membrane calcium-transporting ATPase 1	0.896	0.797	0.826
Carbonyl reductase [NADPH] 3;NADPH-dependent carbonyl reductase 3	0.896	1.061	1.298
DNA-directed RNA polymerase II subunit H;DNA-directed RNA polymerases I, II, and III 17.1 kDa polypeptide;DNA-directed RNA polymerases I, II, and III subunit RPABC3;RPB17;RPB8 homolog;Putative uncharacterized protein POLR2H	0.892	1.439	1.271
70 kDa lamin;Lamin-A/C;Renal carcinoma antigen NY-REN-32;Lamin A/C;Progerin;Rhabdomyosarcoma antigen MU-RMS-40.12	0.888	0.914	0.853
22 kDa neuronal tissue-enriched acidic protein;Brain acid soluble protein 1;Neuronal axonal membrane protein NAP-22	0.888	0.915	0.984
ADP-ribosylation factor-like protein 3	0.884	1.055	1.005
Acetylglucosamine phosphomutase;N-acetylglucosamine-phosphate mutase;Phosphoacetylglucosamine mutase;Phosphoglucomutase-3	0.884	0.712	0.632
Cytochrome c oxidase polypeptide VIIc;Cytochrome c oxidase subunit 7C, mitochondrial	0.879	0.913	0.839

Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 67 kDa subunit;Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1;Ribophorin I;Ribophorin-1;cDNA FLJ50809, highly similar to Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 67 kDa subunit (EC 2.4.1.119);cDNA FLJ51908, highly similar to Dolichyl-diphosphooligosaccharide--proteinglycosyltransferase 67 kDa subunit (EC 2.4.1.119)	0.878	0.715	0.974
Cytochrome c oxidase subunit 7A2, mitochondrial;Cytochrome c oxidase subunit VIIa-liver/heart	0.875	0.705	0.933
21 kDa transmembrane-trafficking protein;p24delta;S31III125;Tmp-21-I;Transmembrane emp24 domain-containing protein 10;Transmembrane protein Tmp21	0.875	0.690	0.754
Cell cycle control protein 50A;Transmembrane protein 30A;cDNA FLJ55687, highly similar to Cell cycle control protein 50A	0.868	2.133	1.883
Oxidative stress-responsive 1 protein;Serine/threonine-protein kinase OSR1;Putative uncharacterized protein OXSR1	0.864	1.203	1.066
Acyl-CoA-binding domain-containing protein 3;Golgi complex-associated protein 1;Golgi phosphoprotein 1;Golgi resident protein GCP60;PBR- and PKA-associated protein 7;Peripheral benzodiazepine receptor-associated protein PAP7	0.861	1.201	1.111
Pyruvate carboxylase, mitochondrial;Pyruvic carboxylase;cDNA FLJ60715, highly similar to Pyruvate carboxylase, mitochondrial (EC 6.4.1.1)	0.857	1.001	0.885
Double-stranded RNA-binding protein Staufen homolog 1;Staufen, RNA binding protein, homolog 1 (Drosophila)	0.855	1.316	1.229
Q15149-6	0.851	0.704	1.087
40S ribosomal protein S19	0.848	0.732	0.944
Intestine-specific plastin;Plastin-1	0.845	1.185	1.053
Nidogen-2;Osteonidogen	0.836	1.244	1.104
Gastric cancer antigen Ga19;N-alpha-acetyltransferase 15, NatA auxiliary subunit;NMDA receptor-regulated protein 1;N-terminal acetyltransferase;Protein tubedown-1;Tbdn100	0.833	0.751	0.704
40S ribosomal protein S14	0.832	0.654	0.615
Heterogeneous nuclear ribonucleoprotein L;cDNA FLJ75895, highly similar to Homo sapiens heterogeneous nuclear ribonucleoprotein L (HNRPL), transcript variant 2, mRNA;Putative uncharacterized protein HNRNPL	0.829	0.747	0.663
cDNA FLJ56102, highly similar to Homo sapiens calpastatin (CAST), transcript variant 8, mRNA;Calpain inhibitor;Calpastatin;Sperm BS-17 component;cDNA FLJ56123, highly similar to Calpastatin	0.826	0.782	0.694
Tropomyosin 3;Tropomyosin 3, isoform CRA_b;cDNA FLJ35393 fis, clone SKNSH2000971, highly similar to TROPOMYOSIN, CYTOSKELETAL TYPE	0.826	0.829	0.732
Macropain chain Z;Multicatalytic endopeptidase complex chain Z;Proteasome subunit beta type-7;Proteasome subunit Z;Proteasome (Prosome, macropain) subunit, beta type, 7;cDNA	0.825	0.745	0.685

FLJ60039, highly similar to Proteasome subunit beta type 7 (EC 3.4.25.1);Proteasome (Prosome, macropain) subunit, beta type, 7, isoform CRA_b			
DNA-directed RNA polymerase II 140 kDa polypeptide;DNA-directed RNA polymerase II subunit B;DNA-directed RNA polymerase II subunit RPB2;RNA polymerase II subunit 2;RNA polymerase II subunit B2;Putative uncharacterized protein POLR2B	0.824	1.499	1.425
40S ribosomal protein S7;Putative uncharacterized protein RPS7	0.824	0.813	0.982
Thyroid hormone receptor-associated protein 3;Thyroid hormone receptor-associated protein complex 150 kDa component	0.823	1.280	1.131
Large tumor suppressor homolog 1;Serine/threonine-protein kinase LATS1;WARTS protein kinase;LATS1 protein	0.822	0.768	0.718
11S regulator complex subunit beta;Activator of multicatalytic protease subunit 2;Proteasome activator 28 subunit beta;Proteasome activator complex subunit 2	0.819	1.303	1.157
Leukocyte common antigen;Receptor-type tyrosine-protein phosphatase C;T200	0.819	0.976	0.975
Tight junction protein 2;Tight junction protein ZO-2;Zona occludens protein 2;Zonula occludens protein 2	0.811	0.919	0.836
CDC42 GTPase-activating protein;GTPase-activating protein rhoOGAP;p50-RhoGAP;Rho GTPase-activating protein 1;Rho-related small GTPase protein activator;Rho-type GTPase-activating protein 1	0.805	0.963	1.011
Calcium-activated neutral proteinase 2;Calpain large polypeptide L2;Calpain M-type;Calpain-2 catalytic subunit;Calpain-2 large subunit;Millimolar-calpain	0.804	1.056	1.017
ABP125;ABP130;Protein transport protein Sec31A;SEC31-like protein 1;SEC31-related protein A;Web1-like protein;SEC31A protein	0.798	0.622	0.895
Electron transfer flavoprotein subunit beta	0.797	0.696	0.886
Collapsin response mediator protein 2;Dihydropyrimidinase-related protein 2;N2A3;Unc-33-like phosphoprotein 2	0.796	0.874	1.048
DEAD box protein 27;Probable ATP-dependent RNA helicase DDX27	0.781	1.299	1.171
Copper amine oxidase;HPAO;Membrane primary amine oxidase;Semicarbazide-sensitive amine oxidase;Vascular adhesion protein 1	0.778	0.845	0.888
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 63 kDa subunit;Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 2;RIBIIR;Ribophorin II;Ribophorin-2	0.777	0.766	1.033
Argininosuccinate synthase;Citrulline--aspartate ligase	0.776	0.987	0.880
Glycoprotein 25L2;Transmembrane emp24 domain-containing protein 9	0.772	1.404	1.242
NAP-1-related protein;Nucleosome assembly protein 1-like 1;cDNA FLJ30458 fis, clone BRACE2009421, highly similar to	0.763	1.080	0.954

NUCLEOSOME ASSEMBLY PROTEIN 1-LIKE 1;cDNA FLJ58569, highly similar to Nucleosome assembly protein 1-like 1;cDNA FLJ16112 fis, clone 3NB692001853, highly similar to NUCLEOSOME ASSEMBLY PROTEIN 1-LIKE 1;Nucleosome assembly protein 1-like 1, isoform CRA_c			
TESS;Testin	0.761	1.013	0.896
14 kDa phosphohistidine phosphatase;Phosphohistidine phosphatase 1;Protein janus-A homolog	0.761	0.888	0.888
Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1;Transducin beta chain 1;Guanine nucleotide binding protein (G protein), beta polypeptide 1	0.758	1.071	1.024
130 kDa leucine-rich protein;GP130;Leucine-rich PPR motif-containing protein, mitochondrial	0.756	0.741	0.954
Cellugyrin;Synaptogyrin-2	0.754	0.685	0.839
c-Ki-ras;c-K-ras;GTPase KRas;GTPase KRas, N-terminally processed;Ki-Ras;K-Ras 2	0.750	1.185	1.075
80K-H protein;Glucosidase 2 subunit beta;Glucosidase II subunit beta;Protein kinase C substrate 60.1 kDa protein heavy chain	0.749	0.874	1.036
Haptoglobin;Haptoglobin alpha chain;Haptoglobin beta chain;HP protein	0.746	0.838	1.044
49 kDa TATA box-binding protein-interacting protein;54 kDa erythrocyte cytosolic protein;INO80 complex subunit H;Nuclear matrix protein 238;Pontin 52;RuvB-like 1;TIP49a;TIP60-associated protein 54-alpha	0.742	0.585	0.524
Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1;Peptidyl-prolyl cis-trans isomerase Pin1;Rotamase Pin1	0.741	0.769	0.689
Acid ceramidase;Acid ceramidase subunit alpha;Acid ceramidase subunit beta;Acylsphingosine deacylase;N-acylsphingosine amidohydrolase;Putative 32 kDa heart protein	0.740	1.345	1.559
Annexin A2;Annexin II;Annexin-2;Calpactin I heavy chain;Calpactin-1 heavy chain;Chromobindin-8;Lipocortin II;p36;Placental anticoagulant protein IV;Protein I;Annexin A2 pseudogene 2;Lipocortin II pseudogene;Putative annexin A2-like protein;cDNA FLJ34687 fis, clone MESAN2000620, highly similar to Annexin A2	0.740	0.626	0.871
Translocation protein 1;Translocation protein SEC62	0.735	1.515	1.341
Smu-1 suppressor of mec-8 and unc-52 protein homolog;WD40 repeat-containing protein SMU1;cDNA FLJ54259, highly similar to Smu-1 suppressor of mec-8 and unc-52 protein homolog	0.723	1.363	1.272
UPF0568 protein C14orf166	0.722	0.730	0.998
NADPH--cytochrome P450 reductase	0.721	1.142	1.009
60S ribosomal protein L3;HIV-1 TAR RNA-binding protein B;Putative uncharacterized protein RPL3	0.715	0.561	0.512
Actin-related protein 2/3 complex subunit 5;Arp2/3 complex 16 kDa subunit	0.709	0.599	0.724
Citrate synthase, mitochondrial;Citrate synthase	0.706	0.689	0.932
ATP-dependent 61 kDa nucleolar RNA helicase;DEAD box protein	0.705	0.549	0.971

21;DEAD box protein 56;Probable ATP-dependent RNA helicase DDX56;Putative uncharacterized protein DDX56			
Ribosome maturation protein SBDS;Shwachman-Bodian-Diamond syndrome protein	0.705	0.681	0.704
Selenide, water dikinase 1;Selenium donor protein 1;Selenophosphate synthase 1;cDNA FLJ60186, highly similar to Selenide, water dikinase 1 (EC 2.7.9.3);Selenophosphate synthetase 1;Selenophosphate synthetase 1, isoform CRA_a	0.699	1.296	1.294
eIF-2B GDP-GTP exchange factor subunit epsilon;Translation initiation factor eIF-2B subunit epsilon	0.691	1.419	1.313
C-terminal LIM domain protein 1;Elfin;LIM domain protein CLP-36;PDZ and LIM domain protein 1	0.673	0.568	0.950
Calcyclin;Growth factor-inducible protein 2A9;MLN 4;Prolactin receptor-associated protein;Protein S100-A6;S100 calcium-binding protein A6	0.672	0.522	0.519
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 48 kDa subunit	0.670	0.522	0.898
ICD-M;IDP;Isocitrate dehydrogenase [NADP], mitochondrial;NADP(+)-specific ICDH;Oxalosuccinate decarboxylase	0.669	1.139	1.036
100 kDa coactivator;EBNA2 coactivator p100;p100 co-activator;Staphylococcal nuclease domain-containing protein 1;Tudor domain-containing protein 11	0.665	0.600	0.914
Glutaredoxin-1;Thioltransferase-1	0.656	1.176	1.038
Microsomal triglyceride transfer protein large subunit	0.639	1.044	1.044
Beta-coat protein;Coatomer subunit beta	0.638	0.631	0.946
60S ribosomal protein L23a;Putative uncharacterized protein RPL23A;Ribosomal protein L23a, isoform CRA_a	0.634	0.711	0.945
14-3-3 protein eta;Protein AS1;Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, eta polypeptide	0.626	1.253	1.129
Putative uncharacterized protein SUMO1;GAP-modifying protein 1;Sentrin;Small ubiquitin-related modifier 1;SMT3 homolog 3;Ubiquitin-homology domain protein PIC1;Ubiquitin-like protein SMT3C;Ubiquitin-like protein UBL1;SMT3 suppressor of mif two 3 homolog 1 (Yeast), isoform CRA_c;SMT3 suppressor of mif two 3 homolog 1 (Yeast), isoform CRA_b	0.623	0.485	0.694
Complex I-15 kDa;NADH dehydrogenase [ubiquinone] iron-sulfur protein 5;NADH-ubiquinone oxidoreductase 15 kDa subunit	0.622	0.602	0.912
Importin-7;Ran-binding protein 7	0.621	0.754	0.722
Dnm1p/Vps1p-like protein;Dynamin family member proline-rich carboxyl-terminal domain less;Dynamin-1-like protein;Dynamin-like protein;Dynamin-like protein 4;Dynamin-like protein IV;Dynamin-related protein 1;cDNA FLJ59504, highly similar to Dynamin-1-like protein (EC 3.6.5.5)	0.614	1.328	1.172
Activated-platelet protein 1;Inducible poly(A)-binding protein;Polyadenylate-binding protein 4;Poly(A) binding protein,	0.611	0.710	0.994

cytoplasmic 4 (Inducible form);Poly(A) binding protein, cytoplasmic 4 (Inducible form), isoform CRA_e			
Succinyl-CoA ligase [GDP-forming] subunit alpha, mitochondrial;Succinyl-CoA synthetase subunit alpha	0.604	0.498	0.906
27 kDa prosomal protein;Macropain iota chain;Multicatalytic endopeptidase complex iota chain;Proteasome iota chain;Proteasome subunit alpha type-6;cDNA FLJ51729, highly similar to Proteasome subunit alpha type 6 (EC 3.4.25.1);Proteasome (Prosome, macropain) subunit, alpha type, 6, isoform CRA_a;cDNA FLJ52022, highly similar to Proteasome subunit alpha type 6 (EC 3.4.25.1);cDNA, FLJ79122, highly similar to Proteasome subunit alpha type 6 (EC 3.4.25.1)	0.603	1.148	1.138
Cytosol aminopeptidase;Leucine aminopeptidase 3;Leucyl aminopeptidase;Peptidase S;Proline aminopeptidase;Prolyl aminopeptidase	0.599	0.963	0.857
60S ribosomal protein L14;CAG-ISL 7;cDNA FLJ51325, highly similar to 60S ribosomal protein L14	0.589	0.522	0.744
DNA topoisomerase 1;DNA topoisomerase I	0.575	0.715	0.997
40S ribosomal protein S26;Putative 40S ribosomal protein S26-like 1	0.571	0.777	0.753
Rho GDP-dissociation inhibitor 1;Rho-GDI alpha;cDNA FLJ50748, highly similar to Rho GDP-dissociation inhibitor 1;Putative uncharacterized protein ARHGDIA	0.562	0.438	1.074
Hsc70/Hsp90-organizing protein;Renal carcinoma antigen NY-REN-11;Stress-induced-phosphoprotein 1;Transformation-sensitive protein IEF SSP 3521	0.550	0.715	1.000
Signal sequence receptor subunit delta;Translocon-associated protein subunit delta;Putative uncharacterized protein SSR4	0.548	0.602	0.556
Putative uncharacterized protein MUTED	0.547	0.734	0.906
ATPase family AAA domain-containing protein 3A	0.544	0.612	0.978
Nuclear distribution protein C homolog;Nuclear migration protein nudC	0.540	0.429	0.779
Protein LSM12 homolog	0.530	0.932	1.102
Protein transport protein Sec61 subunit beta	0.528	0.579	0.791
Macropain subunit C5;Multicatalytic endopeptidase complex subunit C5;Proteasome component C5;Proteasome gamma chain;Proteasome subunit beta type-1	0.509	1.172	1.039
Alcohol dehydrogenase 5;Alcohol dehydrogenase class chi chain;Alcohol dehydrogenase class-3;Alcohol dehydrogenase class-III;Glutathione-dependent formaldehyde dehydrogenase;S-(hydroxymethyl)glutathione dehydrogenase	0.508	0.456	0.936
DEAD box protein 47;Probable ATP-dependent RNA helicase DDX47	0.502	0.638	0.755
Macropain delta chain;Multicatalytic endopeptidase complex delta chain;Proteasome delta chain;Proteasome subunit beta type-6;Proteasome subunit Y	0.499	1.013	1.008
Proteasome chain 13;Proteasome component C10-II;Proteasome	0.499	1.147	1.422

subunit beta type-3;Proteasome theta chain			
Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-5	0.493	0.673	1.014
60S ribosomal protein L35	0.480	0.523	0.899
Putative uncharacterized protein DBN1;Developmentally-regulated brain protein;Drebrin	0.468	0.770	0.774
tRNA pseudouridine synthase A;tRNA pseudouridylate synthase I;tRNA-uridine isomerase I	0.467	0.877	0.963
Protein NDRG2;Protein Syld709613;cDNA FLJ55190, highly similar to Protein NDRG2	0.467	1.647	1.485
AKAP 120-like protein;A-kinase anchor protein 350 kDa;A-kinase anchor protein 450 kDa;A-kinase anchor protein 9;Centrosome- and Golgi-localized PKN-associated protein;Protein hyperion;Protein kinase A-anchoring protein 9;Protein yotiao	0.465	0.785	0.703
Anandamide amidohydrolase 1;Fatty-acid amide hydrolase 1;Oleamide hydrolase 1	0.462	0.612	1.046
Acyl-coenzyme A thioesterase 13;Thioesterase superfamily member 2	0.461	1.134	1.032
40S ribosomal protein S21;RPS21 protein;Ribosomal protein S21;Ribosomal protein S21, isoform CRA_e	0.456	0.594	0.980
NSFL1 cofactor p47;p97 cofactor p47;UBX domain-containing protein 2C	0.448	0.600	1.023
Glutaredoxin-3;PKC-interacting cousin of thioredoxin;PKC-theta-interacting protein;Thioredoxin-like protein 2	0.445	0.808	0.971
Charcot-Leyden crystal protein;CLC;Eosinophil lysophospholipase;Galectin-10;Lysolecithin acylhydrolase	0.445	1.579	1.477
LanC-like protein 2;Testis-specific adriamycin sensitivity protein	0.441	1.599	1.423
Dimethylallyltranstransferase;Farnesyl diphosphate synthase;Farnesyl pyrophosphate synthase;Geranyltranstransferase	0.441	0.469	0.821
26S protease regulatory subunit 7;26S proteasome AAA-ATPase subunit RPT1;Proteasome 26S subunit ATPase 2;Protein MSS1;cDNA FLJ52353, highly similar to 26S protease regulatory subunit 7	0.436	0.409	0.774
CCT-epsilon;T-complex protein 1 subunit epsilon	0.432	0.768	0.859
Glyceraldehyde-3-phosphate dehydrogenase	0.428	0.605	0.994
Ras-related protein Rab-6A	0.427	0.874	1.046
Nucleosome assembly protein 1-like 4b;Putative uncharacterized protein NAP1L4;Nucleosome assembly protein 1-like 4;Nucleosome assembly protein 2	0.426	0.331	0.996
Myosin-VI;Unconventional myosin-6	0.422	0.789	0.865
26S proteasome non-ATPase regulatory subunit 3;26S proteasome regulatory subunit RPN3;26S proteasome regulatory subunit S3;Proteasome subunit p58;cDNA FLJ54148, highly similar to 26S proteasome non-ATPase regulatory subunit 3	0.420	0.562	0.992
Placental ribonuclease inhibitor;Ribonuclease inhibitor;Ribonuclease/angiogenin inhibitor 1	0.419	0.470	0.883

30 kDa prosomal protein;Macropain subunit C2;Multicatalytic endopeptidase complex subunit C2;Proteasome component C2;Proteasome nu chain;Proteasome subunit alpha type-1;Proteasome subunit alpha type	0.402	0.563	0.994
IMPDH-II;Inosine-5-monophosphate dehydrogenase 2	0.401	0.356	0.652
Bax antagonist selected in saccharomyces 1;Negative regulatory element-binding protein;Protein DBP-5;Protein SON;SON3	0.401	0.560	0.906
AGX-1;AGX-2;Antigen X;Sperm-associated antigen 2;UDP-N-acetylgalactosamine pyrophosphorylase;UDP-N-acetylglucosamine pyrophosphorylase;UDP-N-acetylhexosamine pyrophosphorylase;UDP-N-acteylglucosamine pyrophosphorylase 1	0.400	1.176	1.039
Heat shock 70 kDa protein 4;Heat shock 70-related protein APG-2;HSP70RY	0.390	1.029	1.069
Autocrine motility factor;Glucose-6-phosphate isomerase;Neuroleukin;Phosphoglucose isomerase;Phosphohexose isomerase;Sperm antigen 36	0.383	0.317	0.849
ATP-dependent helicase SMARCA2;BRG1-associated factor 190B;Probable global transcription activator SNF2L2;Protein brahma homolog;SNF2-alpha;SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 2	0.379	0.669	1.095
PEP11 homolog;Vacuolar protein sorting-associated protein 29;Vesicle protein sorting 29	0.366	1.199	1.100
Deubiquitinating enzyme 7;Herpesvirus-associated ubiquitin-specific protease;Ubiquitin carboxyl-terminal hydrolase 7;Ubiquitin thioesterase 7;Ubiquitin-specific-processing protease 7;Ubiquitin carboxyl-terminal hydrolase	0.365	0.730	0.991
Butyrate-induced protein 1;Protein tyrosine phosphatase-like protein PTPLAD1;Protein-tyrosine phosphatase-like A domain-containing protein 1;cDNA FLJ54138, highly similar to Homo sapiens butyrate-induced transcript 1 (HSPC121), mRNA	0.365	0.301	0.736
Amine oxidase [flavin-containing] B;Monoamine oxidase type B;cDNA FLJ51821, highly similar to Amine oxidase (flavin-containing) B (EC 1.4.3.4);cDNA FLJ52418, highly similar to Amine oxidase (flavin-containing) B (EC 1.4.3.4)	0.364	1.217	1.126
Cell proliferation-inducing gene 21 protein;Guanine nucleotide-binding protein subunit beta-2-like 1;Guanine nucleotide-binding protein subunit beta-like protein 12.3;Human lung cancer oncogene 7 protein;Receptor for activated C kinase;Receptor of activated protein kinase C 1	0.361	0.644	0.979
CAAX farnesyltransferase subunit alpha;FTase-alpha;Protein farnesyltransferase/geranylgeranyltransferase type-1 subunit alpha;Ras proteins prenyltransferase subunit alpha;Type I protein geranyl-geranyltransferase subunit alpha	0.356	1.417	1.499
1F5 antigen;20 kDa homologous restriction factor;CD59 glycoprotein;MAC-inhibitory protein;MEM43 antigen;Membrane attack complex inhibition factor;Membrane inhibitor of reactive	0.354	1.131	1.368

lysis;Protectin			
Aldehyde dehydrogenase family 1 member A1;Aldehyde dehydrogenase, cytosolic;ALDH-E1;ALHDII;Retinal dehydrogenase 1	0.348	0.678	1.002
Glycine hydroxymethyltransferase;Serine hydroxymethyltransferase, mitochondrial;Serine methylase;cDNA FLJ58585, highly similar to Serine hydroxymethyltransferase, mitochondrial (EC 2.1.2.1);Serine hydroxymethyltransferase 2 (Mitochondrial), isoform CRA_h	0.345	0.677	0.978
Macropain zeta chain;Multicatalytic endopeptidase complex zeta chain;Proteasome subunit alpha type-5;Proteasome zeta chain;cDNA FLJ52182, highly similar to Proteasome subunit alpha type 5 (EC 3.4.25.1);Proteasome (Prosome, macropain) subunit, alpha type, 5, isoform CRA_c	0.339	0.954	0.900
Methionine--tRNA ligase;Methionyl-tRNA synthetase, cytoplasmic;Putative uncharacterized protein MARS;cDNA FLJ16674 fis, clone THYMU3008136, highly similar to Methionyl-tRNA synthetase (EC 6.1.1.10)	0.325	0.537	0.977
Importin-9;Ran-binding protein 9	0.320	0.530	0.801
Pre-mRNA-splicing factor SRP75;Splicing factor, arginine/serine-rich 4;SRP001LB	0.310	1.096	1.336
Anterior gradient protein 2 homolog;HPC8;Secreted cement gland protein XAG-2 homolog;Putative uncharacterized protein AGR2	0.310	1.524	1.365
Coagulation factor XIII A chain;Protein-glutamine gamma-glutamyltransferase A chain;Transglutaminase A chain	0.306	1.379	1.397
DNA replication licensing factor MCM2;Minichromosome maintenance protein 2 homolog;Nuclear protein BM28	0.303	0.555	0.981
Phosphate carrier protein, mitochondrial;Phosphate transport protein;Solute carrier family 25 member 3	0.299	0.569	0.911
CCT-beta;T-complex protein 1 subunit beta	0.291	0.540	0.980
Protein ftsJ homolog 3;Putative rRNA methyltransferase 3;rRNA (uridine-2-O-)-methyltransferase 3	0.286	0.995	1.042
High mobility group-like nuclear protein 2 homolog 1;NHP2-like protein 1;OTK27;SNU13 homolog;U4/U6.U5 tri-snRNP 15.5 kDa protein;NHP2 non-histone chromosome protein 2-like 1 (S. cerevisiae)	0.280	0.676	0.992
CCT-gamma;hTRiC5;T-complex protein 1 subunit gamma	0.278	0.521	0.867
Nuclear matrix protein 200;Pre-mRNA-processing factor 19;PRP19/PSO4 homolog;Senescence evasion factor	0.276	0.291	0.352
Protein mago nashi homolog;cDNA FLJ55283, moderately similar to Protein mago nashi homolog;Mago-nashi homolog, proliferation-associated (Drosophila);Mago-nashi homolog, proliferation-associated (Drosophila), isoform CRA_a	0.272	1.125	1.047
Mannose-6-phosphate isomerase;Phosphohexomutase;Phosphomannose isomerase;Mannose phosphate isomerase isoform	0.261	0.847	0.931

Carnitine/acylcarnitine translocase;Mitochondrial carnitine/acylcarnitine carrier protein;Solute carrier family 25 member 20;cDNA FLJ53016, highly similar to Mitochondrial carnitine/acylcarnitine carrier protein	0.254	1.064	1.024
Caspase-1;Caspase-1 subunit p10;Caspase-1 subunit p20;Interleukin-1 beta convertase;Interleukin-1 beta-converting enzyme;p45;cDNA FLJ59442, highly similar to Caspase-1 (EC 3.4.22.36)	0.252	1.158	1.068
40S ribosomal protein S4, X isoform;SCR10;Single copy abundant mRNA protein	0.252	0.522	0.958
Dipeptidyl aminopeptidase II;Dipeptidyl peptidase 2;Dipeptidyl peptidase 7;Dipeptidyl peptidase II;Quiescent cell proline dipeptidase	0.248	1.042	0.922
Protein mago nashi homolog 2;Putative uncharacterized protein MAGOHB	0.240	1.337	1.259
Brain-type aldolase;Fructose-bisphosphate aldolase C;Fructose-bisphosphate aldolase;Putative uncharacterized protein ALDOC	0.236	0.954	0.892
Ras-related protein Rab-1A;YPT1-related protein;cDNA FLJ57768, highly similar to Ras-related protein Rab-1A	0.230	0.572	0.900
Brush border myosin I;Myosin I heavy chain;Myosin-Ia	0.222	1.238	1.241
CDC21 homolog;DNA replication licensing factor MCM4;P1-CDC21	0.220	0.172	0.646
28 kDa heat shock protein;Estrogen-regulated 24 kDa protein;Heat shock 27 kDa protein;Heat shock protein beta-1;Stress-responsive protein 27;cDNA FLJ52243, highly similar to Heat-shock protein beta-1	0.208	0.621	0.941
ATP-dependent RNA helicase DDX19A;DDX19-like protein;DEAD box protein 19A;ATP-dependent RNA helicase DDX19B;DEAD box protein 19B;DEAD box RNA helicase DEAD5;cDNA FLJ52463, highly similar to ATP-dependent RNA helicase DDX19A (EC 3.6.1.-)	0.207	0.443	0.447
Double-stranded RNA-binding protein 76;Interleukin enhancer-binding factor 3;M-phase phosphoprotein 4;Nuclear factor associated with dsRNA;Nuclear factor of activated T-cells 90 kDa;Translational control protein 80;Putative uncharacterized protein ILF3;cDNA FLJ58801, highly similar to Interleukin enhancer-binding factor 3	0.202	0.696	0.934
PHD finger protein 6;PHD-like zinc finger protein;cDNA FLJ60207, highly similar to PHD finger protein 6;PHD finger protein 6, isoform CRA_d	0.201	1.197	1.485
Glycosyltransferase 25 family member 1;Hydroxylysine galactosyltransferase 1;Procollagen galactosyltransferase 1	0.197	0.183	1.063
D-fructose-6-phosphate amidotransferase 2;Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] 2;Glutamine:fructose 6 phosphate amidotransferase 2;Hexosephosphate aminotransferase 2	0.197	1.720	1.634
CCT-delta;Stimulator of TAR RNA-binding;T-complex protein 1	0.189	0.562	0.972

subunit delta			
5-3 exoribonuclease 2;DHM1-like protein;cDNA FLJ55645, highly similar to 5-3 exoribonuclease 2 (EC 3.1.11.-)	0.187	0.495	0.838
High mobility group protein 2a;High mobility group protein 4;High mobility group protein B3	0.182	0.609	1.217
CCT-theta;Renal carcinoma antigen NY-REN-15;T-complex protein 1 subunit theta;cDNA FLJ53379, highly similar to T-complex protein 1 subunit theta;cDNA FLJ59382, highly similar to T-complex protein 1 subunit theta	0.181	0.581	0.962
CD63 antigen;Granulophysin;Lysosomal-associated membrane protein 3;Melanoma-associated antigen ME491;Ocular melanoma-associated antigen;OMA81H;Tetraspanin-30;Putative uncharacterized protein CD63;Lysosome-associated membrane protein-3 variant	0.180	0.953	0.884
Sideroflexin-1;Tricarboxylate carrier protein	0.178	0.600	0.965
Phosphopantothenate--cysteine ligase;Phosphopantothenoylcysteine synthetase	0.178	0.853	1.016
CCT-alpha;T-complex protein 1 subunit alpha	0.175	0.864	0.990
Aldehyde dehydrogenase family 18 member A1;Delta-1-pyrroline-5-carboxylate synthase;Gamma-glutamyl kinase;Gamma-glutamyl phosphate reductase;Glutamate 5-kinase;Glutamate-5-semialdehyde dehydrogenase;Glutamyl-gamma-semialdehyde dehydrogenase	0.174	0.757	0.868
Alpha-II spectrin;Fodrin alpha chain;Spectrin alpha chain, brain;Spectrin, non-erythroid alpha chain;Putative uncharacterized protein SPTAN1;cDNA FLJ59116, highly similar to Spectrin alpha chain, brain	0.173	0.562	0.960
Delta(3,5)-Delta(2,4)-dienoyl-CoA isomerase, mitochondrial	0.167	0.512	0.972
Farnesyl-diphosphate farnesyltransferase;FPP:FPP farnesyltransferase;Squalene synthase;cDNA FLJ50548, highly similar to Squalene synthetase (EC 2.5.1.21);cDNA FLJ50447, highly similar to Squalene synthetase (EC 2.5.1.21);cDNA, FLJ78892, highly similar to Squalene synthetase (EC 2.5.1.21);cDNA, FLJ79250, highly similar to Squalene synthetase (EC 2.5.1.21);cDNA, FLJ79430, highly similar to Squalene synthetase (EC 2.5.1.21);cDNA FLJ50660, highly similar to Squalene synthetase (EC 2.5.1.21);cDNA, FLJ79433, highly similar to Squalene synthetase (EC 2.5.1.21);cDNA FLJ33164 fis, clone UTERU2000542, highly similar to Squalene synthetase (EC 2.5.1.21)	0.167	0.280	0.600
DEAH box protein 30;Putative ATP-dependent RNA helicase DHX30	0.162	0.794	1.002
Bifunctional coenzyme A synthase;Dephospho-CoA kinase;Dephospho-CoA pyrophosphorylase;Dephosphocoenzyme A kinase;NBP;Pantetheine-phosphate adenylyltransferase;Phosphopantetheine adenylyltransferase;POV-2	0.162	1.139	1.115

DNA-dependent protein kinase catalytic subunit;DNPK1;p460	0.161	0.579	0.967
CDC46 homolog;DNA replication licensing factor MCM5;P1-CDC46;MCM5 minichromosome maintenance deficient 5, cell division cycle 46 (S. cerevisiae), isoform CRA_c;Minichromosome maintenance complex component 5	0.159	0.349	0.925
Protein A1S9;Ubiquitin-activating enzyme E1;Ubiquitin-like modifier-activating enzyme 1;cDNA FLJ54582, highly similar to Ubiquitin-activating enzyme E1	0.156	0.683	0.940
CNDP dipeptidase 2;Cytosolic non-specific dipeptidase;Glutamate carboxypeptidase-like protein 1;Peptidase A	0.155	1.122	0.999
Ras-related protein Rab-5B;cDNA FLJ60627, highly similar to Ras-related protein Rab-5B;RAB5B, member RAS oncogene family, isoform CRA_b	0.152	0.305	0.369
358 kDa nucleoporin;E3 SUMO-protein ligase RanBP2;Nuclear pore complex protein Nup358;Nucleoporin Nup358;p270;Ran-binding protein 2	0.152	0.806	0.988
Paraspeckle protein 2;RNA-binding motif protein 14;RNA-binding protein 14;RRM-containing coactivator activator/modulator;Synaptotagmin-interacting protein	0.142	0.762	0.779
11-zinc finger protein;CCCTC-binding factor;CTCFL paralog;Transcriptional repressor CTCF;Putative uncharacterized protein CTCF	0.131	0.524	1.286
6-phosphofructokinase, muscle type;Phosphofructo-1-kinase isozyme A;Phosphofructokinase 1;Phosphohexokinase	0.124	0.812	1.019
Cyclin-A/CDK2-associated protein p19;Organ of Corti protein 2;Organ of Corti protein II;p19A;p19skp1;RNA polymerase II elongation factor-like protein;SIII;S-phase kinase-associated protein 1;Transcription elongation factor B	0.123	0.644	0.779
DRG family-regulatory protein 1;Likely ortholog of mouse immediate early response erythropoietin 4;Zinc finger CCCH domain-containing protein 15	0.121	0.869	0.890
39S ribosomal protein L53, mitochondrial	0.119	0.960	0.919
Met-induced mitochondrial protein;Mitochondrial carrier homolog 2	0.116	0.476	0.951
Protein H105e3;Sterol-4-alpha-carboxylate 3-dehydrogenase, decarboxylating;Putative uncharacterized protein NSDHL	0.105	1.222	1.087
Bw-45;HLA class I histocompatibility antigen, B-45 alpha chain;MHC class I antigen B*45	0.095	0.796	0.733
Hsc70-interacting protein;Progesterone receptor-associated p48 protein;Protein FAM10A1;Putative tumor suppressor ST13;Renal carcinoma antigen NY-REN-33;Suppression of tumorigenicity 13 protein;ST13 protein;Putative protein FAM10A5;Putative protein FAM10A4	0.095	0.714	0.961
BRCA1-A complex subunit MERIT40;Mediator of RAP80 interactions and targeting subunit of 40 kDa;New component of the BRCA1-A complex	0.091	0.275	0.504

Ezrin-radixin-moesin-binding phosphoprotein 50;Na(+)/H(+) exchange regulatory cofactor NHE-RF1;Regulatory cofactor of Na(+)/H(+) exchanger;Sodium-hydrogen exchanger regulatory factor 1;Solute carrier family 9 isoform A3 regulatory factor 1	0.088	0.557	0.968
Valyl-tRNA synthetase;Protein G7a;Valine--tRNA ligase	0.084	0.710	0.910
HCV F-transactivated protein 2;Up-regulated during skeletal muscle growth protein 5	0.083	0.468	0.994
Adenylate cyclase-stimulating G alpha protein;Extra large alphas protein;Guanine nucleotide-binding protein G(s) subunit alpha isoforms XLas;Putative uncharacterized protein GNAS;Guanine nucleotide-binding protein G(s) subunit alpha isoforms short	0.073	1.209	1.091
CCT-eta;HIV-1 Nef-interacting protein;T-complex protein 1 subunit eta;Putative uncharacterized protein CCT7;cDNA FLJ59454, highly similar to T-complex protein 1 subunit eta;Chaperonin containing TCP1, subunit 7 (Eta), isoform CRA_a	0.071	0.574	0.934
Collagen alpha-1(VI) chain;cDNA FLJ61362, highly similar to Collagen alpha-1(VI) chain	0.068	1.608	1.496
ARF-binding protein 1;E3 ubiquitin-protein ligase HUWE1;HECT, UBA and WWE domain-containing protein 1;Homologous to E6AP carboxyl terminus homologous protein 9;Large structure of UREB1;Mcl-1 ubiquitin ligase E3;Upstream regulatory element-binding protein 1	0.066	0.661	0.963
Alpha-1-acid glycoprotein 2;Orosomucoid-2	0.060	1.836	1.804
Protein NipSnap homolog 3A;Protein NipSnap homolog 4;Target for Salmonella secreted protein C	0.060	0.931	1.090
Antioxidant protein 1;HBC189;Peroxiredoxin III;Peroxiredoxin-3;Protein MER5 homolog;Thioredoxin-dependent peroxide reductase, mitochondrial	0.058	0.578	0.972
Aspartate carbamoyltransferase;CAD protein;Dihydroorotase;Glutamine-dependent carbamoyl-phosphate synthase	0.049	0.820	0.923
Nuclear mitotic apparatus protein 1;SP-H antigen	0.045	0.486	0.961
50 kDa nucleoporin;Nuclear pore complex protein Nup50;Nuclear pore-associated protein 60 kDa-like;Nucleoporin Nup50	0.043	0.036	0.745
[Acyl-carrier-protein] S-acetyltransferase;[Acyl-carrier-protein] S-malonyltransferase;3-hydroxypalmitoyl-[acyl-carrier-protein] dehydratase;3-oxoacyl-[acyl-carrier-protein] reductase;3-oxoacyl-[acyl-carrier-protein] synthase;Enoyl-[acyl-carrier-protein] reductase;Fatty acid synthase;Oleoyl-[acyl-carrier-protein] hydrolase	0.035	0.582	0.581
DNA mismatch repair protein Msh6;G/T mismatch-binding protein;MutS-alpha 160 kDa subunit;cDNA FLJ55677, highly similar to DNA mismatch repair protein MSH6	0.029	0.674	0.714
Heterogeneous nuclear ribonucleoprotein H;Heterogeneous nuclear ribonucleoprotein H, N-terminally processed	0.025	0.583	0.972
Protein-tyrosine phosphatase 1D;Protein-tyrosine phosphatase	0.018	0.430	0.913

2C;SH-PTP2;SH-PTP3;Tyrosine-protein phosphatase non-receptor type 11			
ATP-dependent RNA helicase A;DEAH box protein 9;Nuclear DNA helicase II	0.006	0.562	0.967
Cysteine dioxygenase type 1;Cysteine dioxygenase type I	0.005	1.364	1.209

Appendix XXV: List of differentially expressed mitochondrial proteins and their variable importance in projection scores (VIP) derived from the calculated PLS-DA model.

Variable	Comp 1 VIP	Comp 2 VIP	Comp 3 VIP
Complex I-PDSW;NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 10;NADH-ubiquinone oxidoreductase PDSW subunit;NADH dehydrogenase (Ubiquinone) 1 beta subcomplex, 10, 22kDa, isoform CRA_a;NDUFB10 protein	1.834	1.448	1.178
Complex I-75kD;NADH-ubiquinone oxidoreductase 75 kDa subunit, mitochondrial;cDNA FLJ60586, highly similar to NADH-ubiquinone oxidoreductase 75 kDa subunit, mitochondrial (EC 1.6.5.3)	1.788	1.323	1.100
Isovaleryl-CoA dehydrogenase, mitochondrial;cDNA FLJ16602 fis, clone TESTI4007816, highly similar to Isovaleryl-CoA dehydrogenase, mitochondrial (EC 1.3.99.10);Isovaleryl Coenzyme A dehydrogenase, isoform CRA_b	1.748	1.559	1.324
Complex I-39kD;NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 9, mitochondrial;NADH-ubiquinone oxidoreductase 39 kDa subunit	1.706	1.331	1.100
Cytochrome c oxidase polypeptide Vb;Cytochrome c oxidase subunit 5B, mitochondrial	1.666	1.237	1.006
Complex III subunit 5;Complex III subunit IX;Cytochrome b-c1 complex subunit 11;Cytochrome b-c1 complex subunit 5;Cytochrome b-c1 complex subunit Rieske, mitochondrial;Rieske iron-sulfur protein;Ubiquinol-cytochrome c reductase 8 kDa protein;Ubiquinol-cytochrome c reductase iron-sulfur subunit;Putative cytochrome b-c1 complex subunit Rieske-like protein 1	1.658	1.249	1.038
Complex III subunit 7;Complex III subunit VII;Cytochrome b-c1 complex subunit 7;QP-C;Ubiquinol-cytochrome c reductase complex 14 kDa protein;cDNA FLJ52271, moderately similar to Ubiquinol-cytochrome c reductase complex 14 kDa protein (EC 1.10.2.2)	1.625	1.220	1.007
Complex III subunit 2;Core protein II;Cytochrome b-c1 complex subunit 2, mitochondrial;Ubiquinol-cytochrome-c reductase complex core protein 2	1.551	1.167	1.207
Angiotensin-binding protein;Microsomal endopeptidase;Mitochondrial oligopeptidase M;Neurolysin, mitochondrial;Neurotensin endopeptidase	1.547	1.997	1.627
Rhodanese;Thiosulfate sulfurtransferase	1.504	1.169	0.954
Complex I-ASHI;NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 8, mitochondrial;NADH-ubiquinone oxidoreductase ASHI subunit;NADH dehydrogenase (Ubiquinone) 1 beta subcomplex, 8, 19kDa;NADH dehydrogenase (Ubiquinone) 1 beta subcomplex, 8, 19kDa, isoform CRA_a;cDNA FLJ52503, highly similar to NADH dehydrogenase (ubiquinone) 1 beta subcomplex subunit 8, mitochondrial (EC 1.6.5.3) (EC 1.6.99.3) (NADH-ubiquinone oxidoreductase ASHI subunit) (Complex I-ASHI) (CI-ASHI)	1.489	1.117	1.207
Iron-sulfur subunit of complex II;Succinate dehydrogenase [ubiquinone]	1.486	1.167	1.040

iron-sulfur subunit, mitochondrial			
Complex III subunit 1;Core protein I;Cytochrome b-c1 complex subunit 1, mitochondrial;Ubiquinol-cytochrome-c reductase complex core protein 1	1.474	1.124	0.922
Complex I-B14.5a;NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 7;NADH-ubiquinone oxidoreductase subunit B14.5a	1.470	1.088	1.194
Complex I-23kD;NADH dehydrogenase [ubiquinone] iron-sulfur protein 8, mitochondrial;NADH-ubiquinone oxidoreductase 23 kDa subunit;TYKY subunit	1.443	1.178	0.982
Cytochrome c oxidase polypeptide IV;Cytochrome c oxidase subunit 4 isoform 1, mitochondrial;Cytochrome c oxidase subunit IV isoform 1;COX4I1 protein	1.432	1.063	0.934
Cytochrome c oxidase polypeptide VIc;Cytochrome c oxidase subunit 6C	1.416	1.074	0.929
Glutathione S-transferase kappa 1;Glutathione S-transferase subunit 13;GST 13-13;GST class-kappa;GSTK1-1	1.404	1.172	0.975
Cytochrome c oxidase polypeptide II;Cytochrome c oxidase subunit 2	1.375	1.120	0.943
GTP-specific succinyl-CoA synthetase subunit beta;Succinyl-CoA ligase [GDP-forming] subunit beta, mitochondrial;Succinyl-CoA synthetase beta-G chain	1.363	1.127	0.961
250/210 kDa paraneoplastic pemphigus antigen;Desmoplakin	1.354	1.040	0.867
Complex I-51kD;NADH dehydrogenase [ubiquinone] flavoprotein 1, mitochondrial;NADH dehydrogenase flavoprotein 1;NADH-ubiquinone oxidoreductase 51 kDa subunit;cDNA FLJ57949, highly similar to NADH-ubiquinone oxidoreductase 51 kDa subunit, mitochondrial (EC 1.6.5.3);cDNA, FLJ79021, highly similar to NADH-ubiquinone oxidoreductase 51 kDa subunit, mitochondrial (EC 1.6.5.3)	1.353	1.034	1.066
Alu corepressor 1;Antioxidant enzyme B166;Liver tissue 2D-page spot 71B;Peroxiredoxin V;Peroxiredoxin-5, mitochondrial;Peroxisomal antioxidant enzyme;PLP;Thioredoxin peroxidase PMP20;Thioredoxin reductase;TPx type VI;Putative uncharacterized protein PRDX5	1.336	0.994	0.821
ICD-M;IDP;Isocitrate dehydrogenase [NADP], mitochondrial;NADP(+)-specific ICDH;Oxalosuccinate decarboxylase	1.320	1.210	1.078
Flavoprotein subunit of complex II;Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial	1.292	1.057	1.032
Brain-type aldolase;Fructose-bisphosphate aldolase C;Fructose-bisphosphate aldolase;Putative uncharacterized protein ALDOC	1.289	1.115	0.906
Complex I-13kD-A;NADH dehydrogenase [ubiquinone] iron-sulfur protein 6, mitochondrial;NADH-ubiquinone oxidoreductase 13 kDa-A subunit	1.261	0.935	0.781
Ethylmalonic encephalopathy protein 1;Hepatoma subtracted clone one protein;Protein ETHE1, mitochondrial	1.252	0.951	0.851
Complex I-B15;NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 4;NADH-ubiquinone oxidoreductase B15 subunit;Putative uncharacterized protein NDUFB4	1.241	0.922	0.749
DEAH box protein 30;Putative ATP-dependent RNA helicase DHX30	1.213	1.450	1.220

Cytochrome c oxidase polypeptide VIIc;Cytochrome c oxidase subunit 7C, mitochondrial	1.211	0.897	0.751
Amine oxidase [flavin-containing] A;Monoamine oxidase type A;cDNA FLJ61220, highly similar to Amine oxidase (flavin-containing) A (EC 1.4.3.4)	1.204	0.970	0.866
Carnitine O-palmitoyltransferase 2, mitochondrial;Carnitine palmitoyltransferase II	1.141	0.869	0.871
Amine oxidase [flavin-containing] B;Monoamine oxidase type B;cDNA FLJ51821, highly similar to Amine oxidase (flavin-containing) B (EC 1.4.3.4);cDNA FLJ52418, highly similar to Amine oxidase (flavin-containing) B (EC 1.4.3.4)	1.066	1.289	1.107
Glutaminase kidney isoform, mitochondrial;K-glutaminase;L-glutamine amidohydrolase	1.033	1.317	1.087
Acyl-coenzyme A thioesterase 13;Thioesterase superfamily member 2	1.006	1.393	1.190
Cathepsin D;Cathepsin D heavy chain;Cathepsin D light chain	1.004	0.784	0.652
78 kDa gastrin-binding protein;Long chain 3-hydroxyacyl-CoA dehydrogenase;Long-chain enoyl-CoA hydratase;TP-alpha;Trifunctional enzyme subunit alpha, mitochondrial	0.999	0.742	0.836
Elongation factor Tu, mitochondrial;P43	0.991	0.754	0.929
Enoyl-CoA hydratase 1;Enoyl-CoA hydratase, mitochondrial;Short-chain enoyl-CoA hydratase	0.968	0.842	0.714
Dihydrolipoamide dehydrogenase;Dihydrolipoyl dehydrogenase, mitochondrial;Glycine cleavage system L protein;cDNA FLJ50515, highly similar to Dihydrolipoyl dehydrogenase, mitochondrial (EC 1.8.1.4);Dihydrolipoyl dehydrogenase	0.956	0.716	0.728
Delta(3),delta(2)-enoyl-CoA isomerase;Diazepam-binding inhibitor-related protein 1;Dodecenoyl-CoA isomerase;DRS-1;Hepatocellular carcinoma-associated antigen 88;Peroxisomal 3,2-trans-enoyl-CoA isomerase;Renal carcinoma antigen NY-REN-1;Putative uncharacterized protein PECI	0.945	0.706	0.578
Catalase	0.924	0.881	0.801
3-ketoacyl-CoA thiolase;Acetyl-CoA acyltransferase;Beta-ketothiolase;TP-beta;Trifunctional enzyme subunit beta, mitochondrial;cDNA FLJ56214, highly similar to Trifunctional enzyme subunit beta, mitochondrial;Putative uncharacterized protein HADHB	0.892	0.696	0.981
Aldehyde dehydrogenase family 6 member A1;Methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial	0.882	0.756	0.616
Malic enzyme 2;NAD-dependent malic enzyme, mitochondrial	0.872	1.419	1.325
Outer mitochondrial membrane protein porin 2;Voltage-dependent anion-selective channel protein 2;Voltage-dependent anion channel 2;cDNA FLJ60120, highly similar to Voltage-dependent anion-selective channel protein 2;cDNA, FLJ78818, highly similar to Voltage-dependent anion-selective channel protein 2	0.871	0.649	0.538
Aspartate aminotransferase, mitochondrial;Fatty acid-binding protein;Glutamate oxaloacetate transaminase 2;Plasma membrane-associated fatty acid-binding protein;Transaminase A	0.869	0.668	0.605
Aldehyde dehydrogenase 5;Aldehyde dehydrogenase family 1 member	0.861	0.893	0.729

B1;Aldehyde dehydrogenase X, mitochondrial;cDNA FLJ51238, highly similar to Aldehyde dehydrogenase X, mitochondrial (EC 1.2.1.3)			
Protein NipSnap homolog 1	0.849	0.635	0.541
Calcium-binding mitochondrial carrier protein Aralar2;Citrin;Mitochondrial aspartate glutamate carrier 2;Solute carrier family 25 member 13	0.840	0.650	0.691
Elongation factor Ts, mitochondrial;Elongation factor Ts	0.822	0.893	0.726
Cytosol aminopeptidase;Leucine aminopeptidase 3;Leucyl aminopeptidase;Peptidase S;Proline aminopeptidase;Prolyl aminopeptidase	0.797	0.590	0.848
Outer mitochondrial membrane protein porin 1;Plasmalemmal porin;Porin 31HL;Porin 31HM;Voltage-dependent anion-selective channel protein 1	0.766	0.602	0.533
Superoxide dismutase [Mn], mitochondrial;Superoxide dismutase	0.723	0.581	0.916
Sideroflexin-1;Tricarboxylate carrier protein	0.721	0.705	1.238
Aldehyde dehydrogenase family 18 member A1;Delta-1-pyrroline-5-carboxylate synthase;Gamma-glutamyl kinase;Gamma-glutamyl phosphate reductase;Glutamate 5-kinase;Glutamate-5-semialdehyde dehydrogenase;Glutamyl-gamma-semialdehyde dehydrogenase	0.716	0.559	1.164
39S ribosomal protein L53, mitochondrial	0.705	0.812	0.703
Glycine hydroxymethyltransferase;Serine hydroxymethyltransferase, mitochondrial;Serine methylase;cDNA FLJ58585, highly similar to Serine hydroxymethyltransferase, mitochondrial (EC 2.1.2.1);Serine hydroxymethyltransferase 2 (Mitochondrial), isoform CRA_h	0.687	0.690	1.139
Antioxidant enzyme AOE372;Peroxiredoxin IV;Peroxiredoxin-4;Thioredoxin peroxidase AO372;Thioredoxin-dependent peroxide reductase A0372	0.651	1.017	1.015
Hematopoietic cell-specific LYN substrate 1;Hematopoietic lineage cell-specific protein;LckBP1;p75	0.635	0.872	0.708
Acetoacetyl-CoA thiolase;Acetyl-CoA acetyltransferase, mitochondrial;T2	0.592	0.646	0.573
Complex I-B8;NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 2;NADH-ubiquinone oxidoreductase B8 subunit	0.570	1.106	0.937
Heat shock-related 70 kDa protein 2;cDNA FLJ40505 fis, clone TESTI2045562, highly similar to HEAT SHOCK-RELATED 70 kDa PROTEIN 2	0.551	1.027	0.850
Pyruvate carboxylase, mitochondrial;Pyruvic carboxylase;cDNA FLJ60715, highly similar to Pyruvate carboxylase, mitochondrial (EC 6.4.1.1)	0.542	0.477	0.394
Hydroxysteroid dehydrogenase-like protein 2;cDNA FLJ61200, highly similar to Homo sapiens hydroxysteroid dehydrogenase like 2 (HSDL2), mRNA	0.509	0.403	0.419
PDHE1-A type I;Pyruvate dehydrogenase E1 component subunit alpha, somatic form, mitochondrial	0.498	0.372	0.801
Antioxidant protein 1;HBC189;Peroxiredoxin III;Peroxiredoxin-3;Protein MER5 homolog;Thioredoxin-dependent peroxide reductase, mitochondrial	0.490	0.707	1.120

3-hydroxybutyrate dehydrogenase type 2;Dehydrogenase/reductase SDR family member 6;Oxidoreductase UCPA;R-beta-hydroxybutyrate dehydrogenase	0.476	0.807	0.883
3-5 RNA exonuclease OLD35;PNPase old-35;Polynucleotide phosphorylase 1;Polynucleotide phosphorylase-like protein;Polyribonucleotide nucleotidyltransferase 1, mitochondrial	0.470	1.236	1.035
3-ketoacyl-CoA thiolase, mitochondrial;Acetyl-CoA acyltransferase;Beta-ketothiolase;Mitochondrial 3-oxoacyl-CoA thiolase;T1	0.461	0.677	1.061
Delta(3,5)-Delta(2,4)-dienoyl-CoA isomerase, mitochondrial	0.458	0.638	1.143
Carnitine/acylcarnitine translocase;Mitochondrial carnitine/acylcarnitine carrier protein;Solute carrier family 25 member 20;cDNA FLJ53016, highly similar to Mitochondrial carnitine/acylcarnitine carrier protein	0.453	2.048	1.697
Protein SCO2 homolog, mitochondrial	0.439	1.323	1.107
HCNPpp;Hippocampal cholinergic neurostimulating peptide;Neuropolypeptide h3;Phosphatidylethanolamine-binding protein 1;Prostatic-binding protein;Raf kinase inhibitor protein;cDNA FLJ51535, highly similar to Phosphatidylethanolamine-binding protein 1	0.411	0.517	0.981
Collapsin response mediator protein 2;Dihydropyrimidinase-related protein 2;N2A3;Unc-33-like phosphoprotein 2	0.403	0.309	1.150
28S ribosomal protein S9, mitochondrial	0.356	2.072	1.692
Met-induced mitochondrial protein;Mitochondrial carrier homolog 2	0.353	0.461	1.075
Phosphate carrier protein, mitochondrial;Phosphate transport protein;Solute carrier family 25 member 3	0.353	0.514	1.140
Acyl-CoA-binding domain-containing protein 3;Golgi complex-associated protein 1;Golgi phosphoprotein 1;Golgi resident protein GCP60;PBR- and PKA-associated protein 7;Peripheral benzodiazepine receptor-associated protein PAP7	0.351	1.694	1.801
HCV F-transactivated protein 2;Up-regulated during skeletal muscle growth protein 5	0.344	0.707	1.117
Protein-tyrosine phosphatase 1D;Protein-tyrosine phosphatase 2C;SH-PTP2;SH-PTP3;Tyrosine-protein phosphatase non-receptor type 11	0.303	0.699	1.024
ATP synthase subunit a;F-ATPase protein 6	0.260	0.193	0.279
Complex III subunit 3;Complex III subunit III;Cytochrome b;Cytochrome b-c1 complex subunit 3;Ubiquinol-cytochrome-c reductase complex cytochrome b subunit	0.235	0.489	0.692
Clathrin heavy chain 1;Clathrin heavy chain on chromosome 17	0.226	0.477	0.730
Glutamate dehydrogenase 1, mitochondrial;cDNA FLJ55203, highly similar to Glutamate dehydrogenase 1, mitochondrial (EC 1.4.1.3);cDNA FLJ16138 fis, clone BRALZ2017531, highly similar to Glutamate dehydrogenase 1, mitochondrial (EC 1.4.1.3);Glutamate dehydrogenase 1, isoform CRA_a;Glutamate dehydrogenase 2, mitochondrial	0.216	0.315	0.598
Succinyl-CoA ligase [GDP-forming] subunit alpha, mitochondrial;Succinyl-CoA synthetase subunit alpha	0.189	0.452	1.125
Complex I-15 kDa;NADH dehydrogenase [ubiquinone] iron-sulfur protein 5;NADH-ubiquinone oxidoreductase 15 kDa subunit	0.163	1.344	1.142

Complex I-49kD;NADH dehydrogenase [ubiquinone] iron-sulfur protein 2, mitochondrial;NADH-ubiquinone oxidoreductase 49 kDa subunit;cDNA, FLJ78876, highly similar to NADH-ubiquinone oxidoreductase 49 kDa subunit, mitochondrial (EC 1.6.5.3)	0.137	0.301	1.161
3-hydroxyisobutyryl-CoA hydrolase, mitochondrial;3-hydroxyisobutyryl-coenzyme A hydrolase	0.106	1.378	1.513
Alcohol dehydrogenase 5;Alcohol dehydrogenase class chi chain;Alcohol dehydrogenase class-3;Alcohol dehydrogenase class-III;Glutathione-dependent formaldehyde dehydrogenase;S-(hydroxymethyl)glutathione dehydrogenase	0.093	0.471	1.107
Electron transfer flavoprotein subunit beta	0.091	0.479	1.047
130 kDa leucine-rich protein;GP130;Leucine-rich PPR motif-containing protein, mitochondrial	0.057	0.480	1.135
ATP-specific succinyl-CoA synthetase subunit beta;Renal carcinoma antigen NY-REN-39;Succinyl-CoA ligase [ADP-forming] subunit beta, mitochondrial;Succinyl-CoA synthetase beta-A chain;Succinate-CoA ligase, ADP-forming, beta subunit	0.024	0.865	0.725
Citrate synthase, mitochondrial;Citrate synthase	0.014	0.627	1.020

Curriculum Vitae

Education

PhD in Microbiology and Immunology

- University of Ottawa, ON, Canada.
- September 2008 – February 2015(expected).

Master in Microbiology and Immunology

- Mansoura University, Egypt.
- April 2004 – February 2007.

Bachelor in Pharmaceutical Sciences

- Mansoura University, Egypt.
- September 1996- June 2001.

Laboratory Experience

Ph.D. project

- Department of Biochemistry, Microbiology and Immunology –University of Ottawa, 451 Smyth Road, Ottawa, ON, Canada.
- September 2008 – Present.
- Conducting research into molecular characterization of gut microbiota composition and their contribution to different health disorders.

Master Project

- Department of Microbiology and Immunology, Mansoura University, Mansoura, Egypt.
- April 2004 - February 2007.
- Conducting research into characterization of efflux pumps in clinical isolates of gram negative rods.

Skills

Laboratory Skills

- Construction, optimization and validation of 16S rRNA-based microarray including the design of probes using a combination of softwares such as PhylArray, ARB project and PRIMEROSE program.
- Experience high throughput microbial screening techniques such as microarray, 454-pyrosequencing and Illumina libraries construction.
- Experience performing molecular techniques such as metagenomic DNA, RNA and protein extraction, PCR, QPCR, SDS-PAGE, Western blot, cloning and ELISA.
- Experience working with in-vivo mice colitis models (germ free and gnotobiotic mice).
- Experience basic microbiological techniques such as isolation of clinical isolates, growing and identification of microorganisms via different staining methods and biochemical tests, antibiotic inhibition assays, flourometric characterization of efflux pump, bioscreens and enzyme activity assays.

Analytical Skills

- Experience different bioinformatics tools for analysis of microbial metagenomic data generated by microarrays, 454-pyrosequencing and Illumina sequencing such as Microarray Data Manager (MADAM), Microarray Data Analysis System (MIDAS), and Multiexperiment Viewer (MeV),Ribosomal Database project (RDP), Metagenome Rapid Annotation Using Subsystem Technology (MG-RAST), Quantitative Insights Into Microbial Ecology (QIIME), Galaxy, MEGAN, Interactive Tree of Life (iTOL).
- Statistical analysis and visualization of the generated results via R, XLStat, and GraphPad Prism.

Awards

- 2nd place, annual Biochemistry, Microbiology and Immunology Poster Symposium. (May 2013). Department of Biochemistry, Microbiology and Immunology, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada.
- Best poster view, annual Biochemistry, Microbiology and Immunology Poster Symposium. (May 2013). Department of Biochemistry, Microbiology and Immunology, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada.
- Scholarship from Egyptian Higher Ministry of Education worth 20,000 per year in addition to the tuition fees (September 2008 – August 2012).
- 1st place, annual Biochemistry, Microbiology and Immunology Seminar Symposium. (February 2012). Department of Biochemistry, Microbiology and Immunology, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada.
- Nominated by the graduate studies to represent the program of microbiology at Canadian Student Health Research Forum (CSHRF) poster presentation competition, Winnipeg, Canada. 2012.

Publications

- Caron AZ, He X, Mottawea W, Seifert EL, Jardine K, Dewar-Darch D, Cron GO, Harper ME, Stintzi A, McBurney MW. The SIRT1 deacetylase protects mice against the symptoms of metabolic syndrome. FASEB J. 2014 Mar; 28(3):1306-16.

- Hoang V, Delatolla R, Abujamel T, Mottawea W, Gadbois A, Laflamme E, Stintzi A. Nitrifying moving bed biofilm reactor (MBBR) biofilm and biomass response to long term exposure to 1 °C. Water Res. 2014 Feb 1; 49:215-24.

Conference Presentations

- D. Mack, T. Abujamel, **W. Mottawea**, A. Stintzi. Butyrate producing microbes at the luminal-mucosal interface are reduced in children newly diagnosed with Crohn's disease. Septmeber 2014. 3rd International symposium on Pediatric

Inflammatory Bowel Diseases (PiBD), Rotterdam, The Netherland. Contributed experimental data used in poster presentation given by David Mack.

- **Walid Mottawea**, Hu Zhou, Turki Abujamel, Marcus Mühlbauer, Daniel Figeys, Christian Jobin, David Mack, and Alain Stintzi. A Proinflammatory Host-Microbe Crosstalk in Pediatric Inflammatory Bowel Disease. June 2013. Canadian Society of Microbiologist (CSM) 63rd Conference, Ottawa, ON, Canada. Poster Presentation.
- Turki Abujamel, **Walid Mottawea**, Mehrdad Hajibabaei, Shadi Shokralla, David Mack, and Alain Stintzi. Signature of Colonic Mucosa-associated Butyrate-producing Bacteria in Children. June 2013. Canadian Society of Microbiologist (CSM) 63rd Conference, Ottawa, ON, Canada. Contributed experimental data used in poster presentation given by Turki Abujamel.
- **Walid Mottawea**, Hu Zhou, Turki Abujamel, Marcus Mühlbauer, Daniel Figeys, Christian Jobin, David Mack, and Alain Stintzi. A causal role for the gut microbiota in pediatric inflammatory bowel disease. June 2013. Ottawa Institute of Systems Biology (OISB) Symposium, Mont Tremblant, QC, Canada. Poster Presentation.
- Turki Abujamel, **Walid Mottawea**, Mehrdad Hajibabaei, Shadi Shokralla, David Mack, and Alain Stintzi. June 2013. Diversity of Butyrate-producing Bacteria from the Human Colonic Ecosystem. Ottawa Institute of Systems Biology (OISB) Symposium, Mont Tremblant, QC, Canada. Contributed experimental data used in poster presentation given by Turki Abujamel.
- Marcus Muehlbauer, **Walid Mottawea**, Turki Abujamel, David R. Mack, Alain Stintzi, Christian Jobin. Atopobium parvulum is a predominant member of the adherent microbiome of pediatric IBD patients and promotes colitis in Il10-/- mice. May 2013. Digestive Disease Week (DDW). Orlando, Florida, USA. Contributed experimental data used in poster presentation given by Marcus Muehlbauer.
- **Walid Mottawea**, Hu Zhou, Turki Abujamel, Marcus Mühlbauer, Daniel Figeys, Christian Jobin, David Mack, and Alain Stintzi. Disruption of the Hydrogen

Sulfide Metabolism in Pediatric Inflammatory Bowel Disease. May 2013. Cell Microbiome and Host Health Symposium. Lisbon, Portugal. Poster Presentation.

- Turki Abujamel, **Walid Mottawea**, David Mack, and Alain Stintzi. Characterization of colonic mucosa-associated butyrate-producing bacteria in pediatric inflammatory bowel disease. May 2013. Cell Microbiome and Host Health Symposium. Lisbon, Portugal. Contributed experimental data used in poster presentation given by Turki Abujamel.
- David Mack, **Walid Mottawea**, Turki Abujamel and Alain Stintzi. Colitogenic Microbiota Reside at The Mucosal Luminal Interface of Children With Inflammatory Bowel Disease. 2012 United European Gastroenterology Week. Amsterdam, The Netherlands. Contributed experimental data used in oral presentation given by David Mack.
- **Walid Mottawea**, Turki Abujamel, David Mack and Alain Stintzi. Gut Microbiota and Inflammatory Bowel Disease. 2012 Canadian Student Health Research Forum (CSHRF), Winnipeg, Canada. Poster presentation.
- **Walid Mottawea**, Turki Abujamel, David Mack, Shadi Shokralla, Mehrdad Hajibabaei, and Alain Stintzi. The Normal Gut Microbiota of Adolescent Children. 2012 Keystone Symposia on Molecular and Cellular Biology: The Microbiome, Keystone, Colorado, USA. Poster presentation.
- Turki Abujamel, **Walid Mottawea**, David Mack, Shadi Shokralla, Mehrdad Hajibabaei, and Alain Stintzi. Diversity of Butyrate Producing Bacteria in Children. 2012 Keystone Symposia on Molecular and Cellular Biology: The Microbiome. Contributed experimental data used in poster presentation given by Turki Abujamel.
- **Walid Mottawea**, David Mack and Alain Stintzi. Pediatric Gut Microbiota in Health and Disease. 2011 Ottawa Institute of Systems Biology, Montebello, Quebec, Canada. Poster presentation.

Teaching Experience

- Winter 2010, 2011, 2012 and 2013. Lab demonstrator and teaching assistant for BCH 2333 Biochemistry laboratory course. University of Ottawa, Ottawa, Canada.
- Fall 2009, 2010, 2011, and 2012. Lab demonstrator and teaching assistant for BCH 3356 Molecular Biology laboratory course. University of Ottawa, Ottawa, Canada.
- Winter 2012 and 2013. Teaching assistant for HSS 1100 Microbiology and Immunology course. University of Ottawa, Ottawa, Canada.
- April 2007-September 2008. Assistant lecturer for Microbiology course. Department of Microbiology and Immunology, Mansoura University, Al-Mansoura, Egypt.
- March 2003 – March 2007. Lab demonstrator for Microbiology course. Department of Microbiology and Immunology, Mansoura University, Al-Mansoura, Egypt.